IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant	: Yu et al.	Art Unit	: 1652
Serial No.	: 10/718,986	Examiner	: Tekchand Saidha, Ph.D.
Filed	: November 21, 2003	Conf. No.	: 3664
Title	: BROAD SPECTRUM ANTI-VIRAL THERAPEUTICS AND PROPHYLAXIS		

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPENDIX

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 269, No. 43, Issue of October 28, pp. 27149–27154, 1994
Printed in U.S.A.

Heparan Sulfate Is Essential to Amphiregulin-induced Mitogenic Signaling by the Epidermal Growth Factor Receptor*

(Received for publication, June 20, 1994, and in revised form, July 18, 1994)

Gibbes R. Johnson‡ and Lily Wong

From the Division of Cytokine Biology, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland 20892

Human amphiregulin (AR) is a heparin-binding growth factor which functions by binding to and activating the epidermal growth factor (EGF) receptor tyrosine kinase. AR contains an EGF-like domain (residues 44–84) and a Lys/Arg-rich NH_2-terminal extension (residues 1–43). Synthetic peptides corresponding to residues 8–26, 26–44, and 68–84 of AR were tested for their ability to compete for the binding of AR to immobilized heparin. AR^{8-26} and AR^{68-84} had no significant effect on the binding of AR to heparin, whereas AR^{26-44} bound to heparin and blocked the binding of AR to heparin. Both soluble heparin and heparan sulfate inhibited AR-induced mitogenesis in MCF-10A human mammary epithelial cells with an IC_{50} of 5 and 2 μg/ml, respectively, whereas soluble chondroitin sulfate had only a slight inhibitory effect. When MCF-10A cells were grown in the presence of chlorate, an inhibitor of sulfation, or exposed to the glycosaminoglycan-degrading enzymes heparitinase or heparinase, the ability of AR to evoke mitogenesis in these cells was lost. Chlorate, heparitinase, or heparinase treatment inhibited AR-induced autophosphorylation of tyrosine residues in the EGF receptor. None of these treatments had any significant effect on EGF-triggered mitogenic signaling by the EGF receptor. These results indicate that extracellular heparan sulfate glycosaminoglycan is essential to AR-induced mitogenic signaling by the EGF receptor tyrosine kinase.

Human amphiregulin (AR)[1] is a heparin-binding polypeptide growth regulator which consists of an epidermal growth factor (EGF)-like domain and a very basic NH_2-terminal extension which contains glycosylation sites and putative nuclear localization signals (1–3). AR influences the proliferation of cells by binding to the extracellular domain of the EGF receptor (EGFR) which results in autophosphorylation of the EGFR, activation of the EGFR tyrosine kinase, and rapid tyrosine phosphorylation of a number of cellular substrates including $p185^{erbB2}$ (4). AR stimulates the proliferation of normal and malignant epithelial cells, fibroblasts, and keratinocytes (1–7). *In vivo*, AR is expressed by a large number of normal tissues (8) but appears to be localized exclusively to the epithelium of the human colon (7, 9), stomach (10, 11), breast (12), and pancreas (13). AR has been shown to drive the proliferation of human colon carcinoma cells via an autocrine mechanism (7) and is commonly overexpressed in human cancers of the colon (7, 9, 14, 15), breast (12, 16), stomach (10, 11, 15), and pancreas (13).

Heparin affinity chromatography has been utilized to purify AR from the conditioned medium of human keratinocytes (2) as well as phorbol ester-treated human breast carcinoma cells (3), and 30 μg/ml of soluble heparin has been shown to inhibit the ability of AR to stimulate the growth of keratinocytes (2). In addition to AR, a number of heparin-binding growth factors have been discovered within the last several years which contain an EGF-like domain such as schwannoma-derived growth factor (SDGF) (17), heparin-binding EGF-like growth factor (HB-EGF) (18), betacellulin (19), heregulin (20), and *neu* differentiation factor (21). The bioactivity of AR appears to be mediated exclusively through the EGFR (4, 7), as may be the case for SDGF, HB-EGF, and betacellulin, whereas the action of heregulin and *neu* differentiation factor appears to involve the *erb*B2, *erb*B3, and/or *erb*B4 EGFR-like tyrosine kinases (22–25). Since it is very unlikely that these growth factors would ever encounter heparin *in vivo* (26), the physiological significance of their ability to bind heparin is not clear. Proteoglycans are proteins which contain covalently attached sulfated glycosaminoglycan (GAG), exist on the surface of cells and in the extracellular matrix and are believed to play important roles in a wide range of biological processes which include cell division, morphogenesis and cancer (26–29). One important subset of these molecules is the heparan sulfate (HS) proteoglycan (30–32) whose HS chain is structurally related to heparin, but in general, is sulfated to a lesser degree. HS proteoglycan has been shown to be obligatory for the mitogenic activity of basic and acidic fibroblast growth factor (FGF) (33–37) and vascular endothelial growth factor (38). HS also appears to play an important role in the HB-EGF stimulation of smooth muscle cell migration (39).

Recently, we isolated multiple, structurally distinct forms of AR (3). The predominant ~16.5-kDa forms contained sialic acid-rich complex *N*-linked oligosaccharide, in addition to *O*-linked carbohydrate. However, a non-glycosylated ~9.5-kDa species was also isolated which contained an intact EGF-like core, but had a truncated NH_2-terminal extension. All of these forms bound strongly to heparin and were biologically active, demonstrating that the oligosaccharide moieties and the extreme NH_2-terminal region of AR are not essential to heparin-binding nor bioactivity (3). This previous work also suggested that the ability of AR to bind heparin may be related to its ability to activate the EGFR tyrosine kinase. In this report, we provide strong evidence that extracellular HS GAG is essential to AR-triggered mitogenic signaling by the EGFR.

EXPERIMENTAL PROCEDURES

Purification of AR—Human AR was purified to homogeneity from the conditioned medium of phorbol 12-myristate 13-acetate-treated MCF-7

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Division of Cytokine Biology, CBER, FDA, HFM-511, Bldg. 29A, 8800 Rockville Pike, Bethesda, MD 20892. Tel.: 301-496-9012; Fax: 301-402-1659.

[1] The abbreviations used are: AR, human amphiregulin; EGF, human epidermal growth factor; EGFR, epidermal growth factor receptor; FGF, fibroblast growth factor; HB-EGF, heparin-binding EGF-like growth factor; GAG, glycosaminoglycan; HS, heparan sulfate; TGF-α, transforming growth factor-α; PAGE, polyacrylamide gel electrophoresis; SDGF, schwannoma-derived growth factor.

human breast carcinoma cells by sequential heparin affinity, immunoaffinity, and reverse phase-high performance chromatography as described in Johnson *et al.* (3).

Preparation of AR Peptides—Peptides corresponding to residues 8–26 and 26–44 of AR were synthesized and purified as described previously (7). The peptide corresponding to residues 68–84 of AR was prepared as described in Johnson *et al.* (3).

Binding of AR to Immobilized Heparin—Twenty-five ng of AR and 5 µl of heparin cross-linked to agarose (~1 µg heparin/µl resin; Sigma) were added to 300 µl of 20 mM Hepes, 50 mM NaCl, pH 7.4 (buffer) in the absence or presence of 20 µg of AR peptide. The mixture was rotated end over end for 4 h at 4 °C and centrifuged, and the pellet was washed three times with 1 ml of buffer. The pellet was boiled for 5 min in 20 µl of SDS-polyacrylamide gel electrophoresis (PAGE) sample buffer. SDS-PAGE and Western blotting were performed as described in Johnson *et al.* (3).

Digestion of AR with N-Glycosidase F—One unit of *N*-glycosidase F (Boehringer Mannheim) was added to 100 ng of AR in 300 µl of 20 mM Hepes, 50 mM NaCl, pH 7.4, and incubated for 4 h at 37 °C.

Cell Culture and Mitogenesis Assay—MCF-10A human mammary epithelial cells were cultured and the mitogenesis assay was performed as described previously (3, 4). Briefly, 64 h after the addition of 250 pM AR or EGF (Life Technologies, Inc.), cells in 96-well plates were pulsed for 6 h with [^{3}H]thymidine (2 µCi/well; Amersham Corp.), DNA was harvested, and the incorporation of [^{3}H]thymidine into DNA was quantitated.

EGFR Autophosphorylation Assay—Ligand-induced autophosphorylation of tyrosine residues in the EGFR was measured as described in Johnson *et al.* (4). MCF-10A cells were plated into 100-mm dishes at a density of 785,000 cells per dish and after 2 days of growth were stimulated with 250 pM AR or EGF for 9 min at 37 °C. The EGFR was immunoprecipitated using E7 antiserum directed against the cytoplasmic domain of the human EGFR (40), fractionated in an 8% SDS-PAGE gel, and transferred to a polyvinyl difluoride membrane, and tyrosine-phosphorylated EGFR was detected using biotinylated PY-20 antibody (ICN Biomedicals), streptavidin-horseradish peroxidase conjugate, and enhanced chemiluminescence (Amersham).

Western Blotting Analysis of EGFR—Prior to immunoprecipitation of the EGFR in the autophosphorylation assay, aliquots of total cell crude lysates were taken to evaluate cellular EGFR levels. Proteins were fractionated in an 8% polyacrylamide SDS-PAGE gel and transferred to a polyvinyl difluoride membrane, and EGFR was detected using E7 antiserum (40), the Vectastain ABC Elite kit (Vector Laboratories), and enhanced chemiluminescence.

RESULTS

Residues 26–44 of AR Interact with Heparin—The binding of AR to immobilized heparin has greatly facilitated the purification of AR derived from the conditioned medium of human keratinocytes (2) and human breast carcinoma cells (3). To study the interaction of AR and heparin, a micro-assay was developed in which AR can be bound to a small quantity of heparin that has been cross-linked to agarose (5 µl). After washing the resin, bound AR can be released by boiling the resin in SDS-PAGE sample buffer. The AR is then fractionated in an SDS-PAGE gel and detected by Western blot analysis (Fig. 1). To identify heparin-binding regions in the AR molecule, various synthetic peptides which correspond to distinct regions of AR were tested for the ability to block the binding of AR to immobilized heparin. Peptides which correspond to residues 8–26 and 68–84 had no significant effect on the binding of AR to immobilized heparin (Fig. 1*A*). However, the peptide corresponding to residues 26–44 bound to heparin, as evidenced by the fact that when the Western blot was performed using antibodies directed against residues 26–44 it resulted in a very strong immunopositive streak running down that lane of the gel (Fig. 1*A*). To confirm that AR$^{26-44}$ bound to the same site on the heparin molecule as did AR and thus, could compete for the binding of AR to heparin, the experiment was repeated and AR was detected in the Western blot using an antibody directed against residues 8–26 of AR (AR-Ab3). Since there are two potential *N*-linked glycosylation sites within residues 8–26 of AR, it was necessary to first digest AR with *N*-glycosidase F so

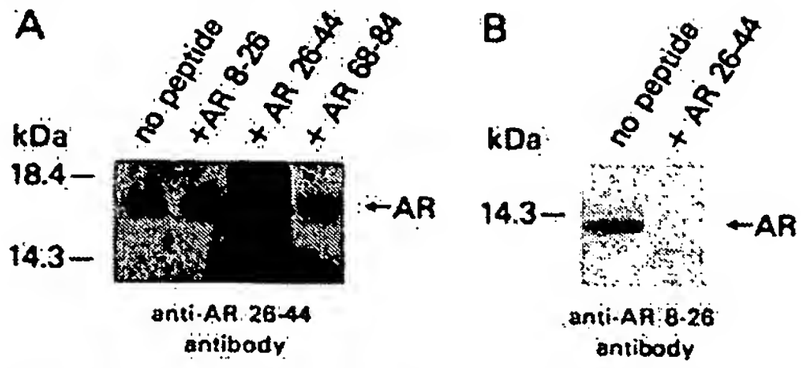


FIG. 1. **Effect of AR peptides on the binding of AR to immobilized heparin.** Twenty-five ng of AR were added to 5 µl of heparin-agarose (~5 µg of immobilized heparin) in 300 µl of 20 mM Hepes, 50 mM NaCl, pH 7.4 (buffer), in the absence or presence of 20 µg of peptide corresponding to residues 8–26 of AR (*AR 8–26*), 26–44 (*AR 26–44*), or 68–84 (*AR 68–84*). After incubation for 4 h at 4 °C, the resin was washed three times with 1 ml of buffer and boiled in SDS-PAGE sample buffer. The sample was fractionated by SDS-PAGE under reducing conditions in a 16% acrylamide gel and transferred to a nitrocellulose membrane. The positions and molecular mass of marker proteins are shown to the left in kilodaltons. In *A*, the nitrocellulose membrane was probed with AR-Ab2 antibodies directed against residues 26–44 of AR (3, 7). In *B*, AR was digested with *N*-glycosidase F prior to performing the heparin-binding assay and the Western blot was probed with AR-Ab3 antibodies directed against residues 8–26 of AR (3). The specificity of the antibodies used was confirmed by performing Western blot analyses using purified control preimmune antibodies (data not shown).

that these antipeptide antibodies could recognize the molecule in Western blot analysis (3). As shown in Fig. 1*B*, the AR$^{26-44}$ peptide completely blocked the binding of AR to heparin. These results demonstrate that residues 26–44 constitute a heparin-binding region in the AR molecule.

Effect of Soluble GAGs on AR-Induced Mitogenesis in Human Mammary Epithelial Cells—Various soluble GAGs were then tested for the ability to affect mitogenesis evoked by AR and EGF in MCF-10A human mammary epithelial cells. MCF-10A is an immortalized nontransformed cell line in which the action of AR is mediated solely by the EGFR tyrosine kinase (4). Previous work from our laboratory has demonstrated that in MCF-10A cells there is an excellent correlation between AR- and EGF-driven increases in cell number and increases in the incorporation of [^{3}H]thymidine into DNA (DNA synthesis) (3, 4). As shown in Fig. 2, both soluble heparin (*Panel A*) and HS (*Panel B*) inhibited mitogenesis induced by 250 pM AR with an IC_{50} of 5 and 2 µg/ml, respectively. However, even at very high concentrations of soluble heparin or HS (100 µg/ml) complete inhibition of AR action was not observed in MCF-10A cells (Fig. 2). Soluble chondroitin sulfate was found to have only a slight inhibitory effect on AR-induced DNA synthesis (Fig. 2*C*) and none of the three GAGs had any significant effect on mitogenesis triggered by EGF (Fig. 2, *A–C*). These results suggested that soluble HS/heparin compete with an extracellular HS-like GAG molecule for binding of AR and that the interaction between this extracellular HS-like molecule and AR is important to the eventual activation of the EGFR.

Chlorate Inhibits AR-triggered Mitogenesis and EGFR Autophosphorylation in MCF-10A Cells—To determine if HS GAG produced by MCF-10A cells is critical to AR-induced mitogenic signaling by the EGFR, two distinct approaches were used. First, cells were grown in the presence of chlorate to specifically interfere with the proper biosynthesis of the sulfated GAG chains. Chlorate is a competitive inhibitor of ATP sulfurylase action because it competes with the recognition of sulfate by the enzyme (41). Growth of cells in the presence of chlorate results in reduced sulfation of GAGs (33). MCF-10A cells which were cultured in 10 mM sodium chlorate lost the ability to respond to exogenous 250 pM AR (Fig. 3*A*). However, chlorate also had a slight inhibitory effect on mitogenesis induced by 250 pM EGF (Fig. 3*B*). To confirm that the inhibition of cell growth caused by chlorate is via the competitive inhibition of sulfation, 5 mM

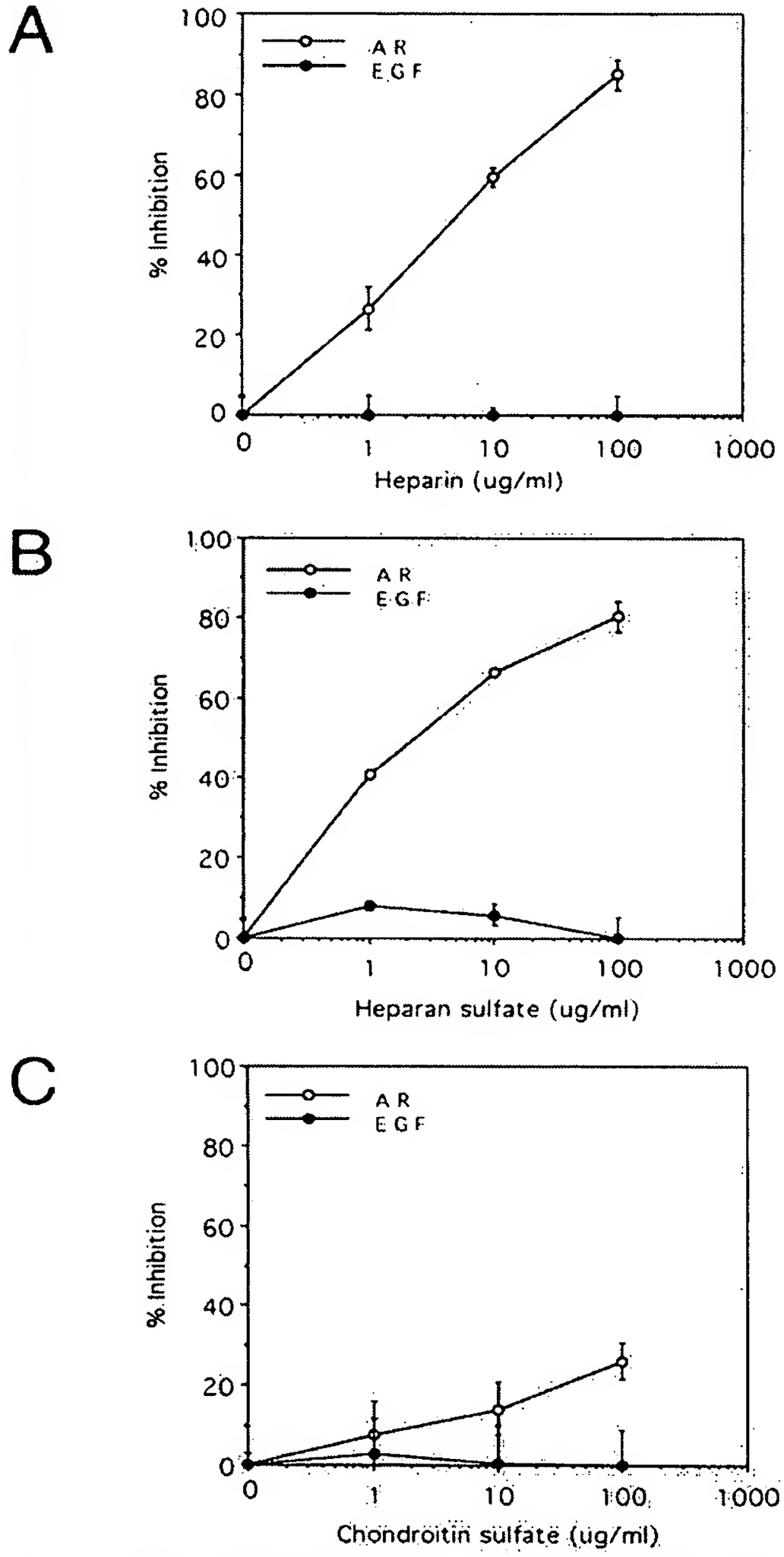


FIG. 2. **Effect of sulfated glycosaminoglycans on AR-induced mitogenesis in human mammary epithelial cells.** MCF-10A cells were cultured as described previously (4), plated into 96-well plates at a density of 2000 cells/well and a mitogenesis assay was performed (3) by adding 250 pM AR or EGF in the absence or presence of various concentrations of soluble heparin (*A*), heparan sulfate (*B*), or chondroitin sulfate (*C*) (Sigma). DNA synthesis was determined by quantifying the incorporation of [^{3}H]thymidine into DNA. Percent inhibition of DNA synthesis for each concentration of glycosaminoglycan was calculated relative to the level of DNA synthesis achieved in the presence of growth factor and the absence of the glycosaminoglycan. Data points represent the mean ± S.E. of experiments performed in triplicate. Values for incorporation of [^{3}H]thymidine into DNA in the absence of glycosaminoglycan for the control (no growth factor), AR and EGF treatment were 8,512, 49,975, and 56,840 cpm/well, respectively. Chondroitin sulfate is a mixture of chondroitin sulfate A and C (Sigma).

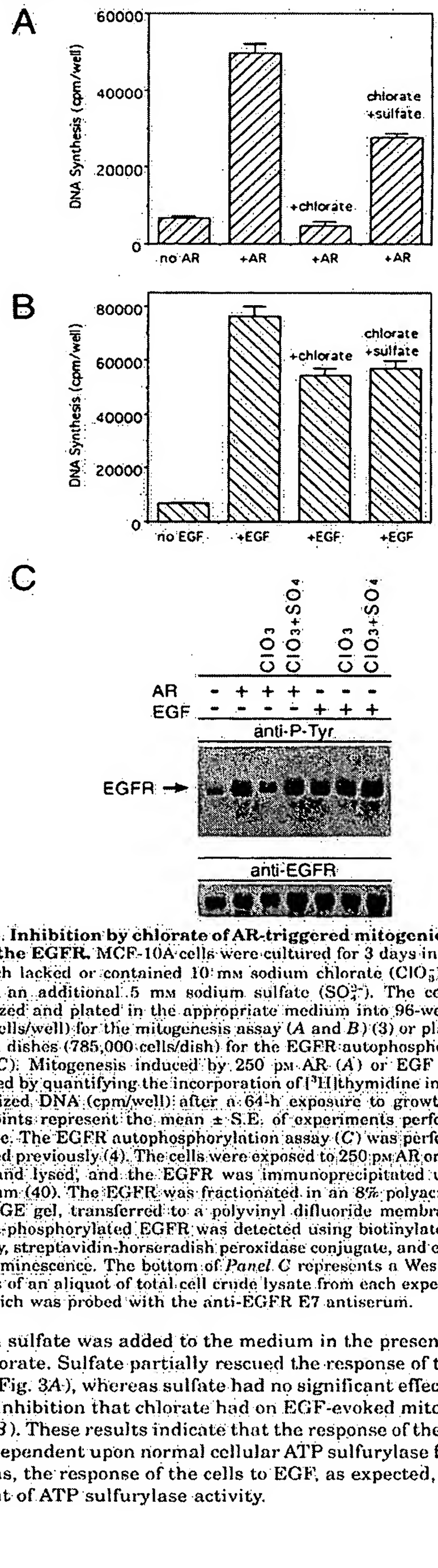


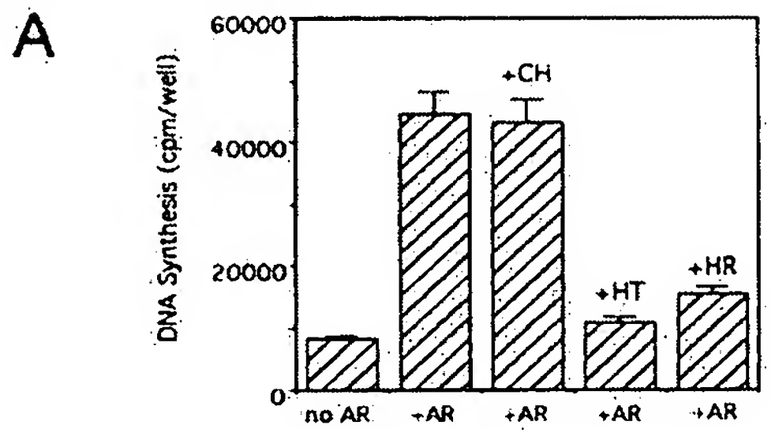
FIG. 3. **Inhibition by chlorate of AR-triggered mitogenic signaling by the EGFR.** MCF-10A cells were cultured for 3 days in medium (4) which lacked or contained 10 mM sodium chlorate (ClO_3^-) with or without an additional 5 mM sodium sulfate (SO_4^{2-}). The cells were trypsinized and plated in the appropriate medium into 96-well plates (2,000 cells/well) for the mitogenesis assay (*A* and *B*) (3) or plated into 100-mm dishes (785,000 cells/dish) for the EGFR autophosphorylation assay (*C*). Mitogenesis induced by 250 pM AR (*A*) or EGF (*B*) was measured by quantifying the incorporation of [^{3}H]thymidine into newly synthesized DNA (cpm/well) after a 64-h exposure to growth factor. Data points represent the mean ± S.E. of experiments performed in triplicate. The EGFR autophosphorylation assay (*C*) was performed as described previously (4). The cells were exposed to 250 pM AR or EGF for 9 min and lysed, and the EGFR was immunoprecipitated using E7 antiserum (40). The EGFR was fractionated in an 8% polyacrylamide SDS-PAGE gel, transferred to a polyvinyl difluoride membrane, and tyrosine-phosphorylated EGFR was detected using biotinylated PY-20 antibody, streptavidin-horseradish peroxidase conjugate, and enhanced chemiluminescence. The bottom of *Panel C* represents a Western blot analysis of an aliquot of total cell crude lysate from each experimental dish which was probed with the anti-EGFR E7 antiserum.

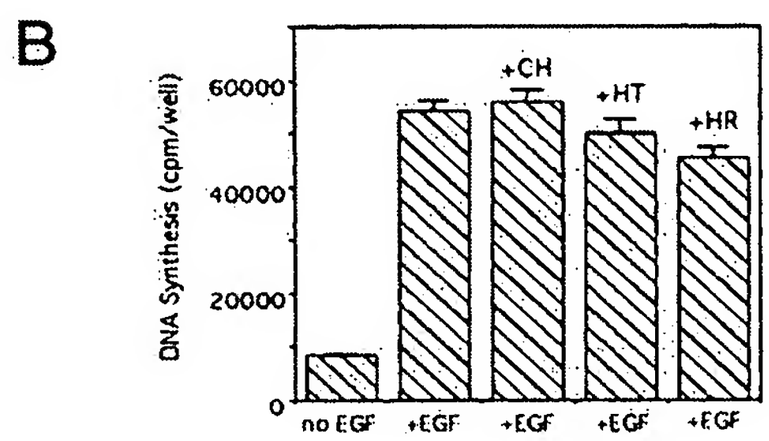
sodium sulfate was added to the medium in the presence of 10 mM chlorate. Sulfate partially rescued the response of the cells to AR (Fig. 3*A*), whereas sulfate had no significant effect on the minor inhibition that chlorate had on EGF-evoked mitogenesis (Fig. 3*B*). These results indicate that the response of the cells to AR is dependent upon normal cellular ATP sulfurylase function whereas, the response of the cells to EGF, as expected, is independent of ATP sulfurylase activity.

The Journal of Biological Chemistry

Mitogenic signaling by the EGFR is believed to occur by ligand-triggered dimerization and autophosphorylation of the EGFR by an intermolecular *trans*-mechanism (42). This activation of the EGFR via autophosphorylation on tyrosine residues results in access of the EGFR tyrosine kinase domain to cytosolic substrates and recruits signaling molecules with Src homology domains (SH2) which can specifically interact with tyrosine-phosphorylated regions of the EGFR. Activation of the EGFR via autophosphorylation appears to be a necessary phenomenon for AR-induced mitogenesis (4). Thus, we investigated the effect that chlorate had on the ability of AR to trigger autophosphorylation of the EGFR. Growth of the cells in 10 mM sodium chlorate significantly reduced the ability of 250 pM AR to drive autophosphorylation of tyrosine residues on the EGFR and the response of the EGFR to AR was completely rescued by 5 mM sodium sulfate (Fig. 3*C*). Neither chlorate nor chlorate plus sulfate had any significant effect on EGF-induced autophosphorylation of the EGFR (Fig. 3*C*). Western blot analysis of cell lysates with anti-EGFR antibodies demonstrated that the treatment of the cells with chlorate had no significant effect on EGFR levels in the MCF-10A cells (bottom of Fig. 3*C*). Therefore, in the case of AR, autophosphorylation of the EGFR and the resultant mitogenic signaling by the EGFR is dependent upon the proper biosyntheis of a sulfated molecule.

Treatment of Cells with Heparitinase or Heparinase Inhibits AR-induced Mitogenic Signaling by the EGFR—The second approach used to study the role that GAGs play in mitogenesis elicited by AR involved the utilization of enzymes which cleave GAG chains at specific sites. Chondroitinase ABC (EC 4.2.2.4) catalyzes the removal of dermatan sulfate and chondroitin sulfate side chains of proteoglycans (43) whereas heparitinase (EC 4.2.2.8) and heparinase (EC 4.2.2.7) cleave distinct sites within HS GAG chains (44). Exposure of MCF-10A cells to either heparitinase or heparinase almost completely blocked the ability of AR to drive mitogenesis, whereas chondroitinase ABC had no effect on this phenomenon (Fig. 4*A*). Heparitinase and heparinase inhibited the AR-stimulated growth of the cells by approximately 93 and 81%, respectively. Conversely, the stimulation of cell division elicited by EGF was not significantly affected by these GAG-degrading enzymes (Fig. 4*B*). Treatment of the cells with heparitinase or heparinase prior to the addition of exogenous AR dramatically inhibited activation of the EGFR as evidenced by the lack of AR-triggered autophosphorylation of tyrosine residues in the EGFR (Fig. 4*C*). In contrast, exposure of the cells to chondroitinase ABC had little or no effect on the response of the EGFR to AR (Fig. 4*C*). These enzymes did not significantly alter the ability of EGF to activate the EGFR nor did they affect EGFR levels in the MCF-10A cells (Fig. 4*C*, *bottom panel*). These results demonstrate that the sulfated GAG which is critical to AR-induced mitogenic signaling by the EGFR is structurally very similar to HS but does not appear to be related to dermatan or chondroitin sulfate.





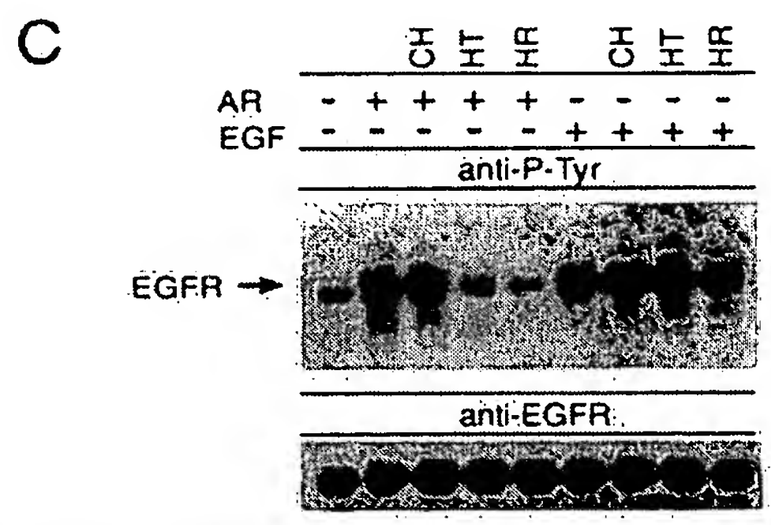


FIG. 4. **Inhibition of AR-induced EGFR mitogenic signaling by heparitinase and heparinase.** MCF-10A cells were plated into 96-well plates at a density of 2,000 cells/well and a mitogenesis assay was performed (3) by adding 250 pM AR (*A*) or EGF (*B*) in the absence or presence of 0.017 unit/ml of chondroitinase ABC (*CH*), heparitinase (*HT*), or heparinase (*HR*) (ICN Biomedicals). Every 12 h, 5 ml of medium with or without 0.0017 unit of the appropriate enzyme were added to each well. Mitogenesis was measured by quantifying the incorporation of [^{3}H]thymidine into newly synthesized DNA (counts/min/well) after a 64-h exposure to growth factor. Data points represent the mean ± S.E. of experiments performed in triplicate. The EGFR autophosphorylation assay (*C*) was performed exactly as described under "Experimental Procedures" except that, prior to exposure to 250 pM AR or EGF, glycosaminoglycan chains were digested by treating the cells for 1 h at 37 °C with 0.02 unit/ml of chondroitinase ABC (*CH*), heparitinase (*HT*), or heparinase (*HR*). The *bottom* of *Panel C* represents a Western blot analysis of an aliquot of total cell crude lysate from each experimental dish which was probed with the anti-EGFR E7 antiserum.

DISCUSSION

AR is a potent stimulator of proliferation in a number of different cell types including normal and malignant epithelial cells, fibroblasts and keratinocytes (1–7). Overexpression of AR has often been observed in human malignancies of the breast, colon, stomach and pancreas (7, 9–16) and in human colon carcinoma cells AR can function as an autocrine growth stimulator (7). However, AR is also expressed by epithelial cells in a number of normal human tissues including the mammary gland (7, 9–13). AR has been shown to act as an autocrine growth factor for normal human mammary epithelial cells (45, 46) and to function as an autocrine growth stimulator in MCF-10A cells when they are transformed by overexpression of activated *ras* or the EGFR-like tyrosine kinase *erb*B2 (47). The work which we have reported here demonstrates that extracellular HS GAG chains play a very important role in AR action in human mammary epithelial cells and are essential to the mitogenic activation of the EGFR which is evoked by AR.

It seems most probable that these HS chains exist on the surface of the cell where they can be covalently linked to an integral membrane protein as in the case of the syndecans (30, 31) or linked to cell surface lipid via glycosyl phosphatidylinositol (PI) as in the case of glypican (32, 48). Treatment of MCF-10A cells with phosphatidylinositol-specific phospholipase C had no effect on AR-induced mitogenesis suggesting that the

HS chains essential to AR functioning are not phosphatidylinositol-anchored.[2] It has been shown that a high affinity receptor for acidic FGF contains covalently attached HS (37). However, even though there are 4 potential GAG attachment sites in the extracellular domain of the EGFR (49) treatment of the cells with the GAG-degrading enzymes had no apparent effect on the migration of the EGFR under conditions of reducing SDS-PAGE (Fig. 4*C*, *bottom panel*). Thus, the HS chains which are critical to AR-triggered mitogenic signaling by the EGFR do not appear to be covalently linked to the EGFR itself, but may be attached to one or more syndecan-like proteoglycans.

On formalin-fixed A431 cells and immobilized plasma membrane preparations, AR is significantly less effective than EGF, at competing for the binding of ^{125}I-EGF to the EGFR (1). Conversely, AR is as potent as EGF in stimulating mitogenesis in MCF-10A cells (4, 47). One interpretation of these findings is that on living cells, HS chains stabilize a mitogenic signaling complex between AR and the EGFR. In the case of basic FGF, soluble heparin can substitute for the presence of the HS proteoglycan and can reconstitute the biological action of basic FGF in cells which lack the requisite HS proteoglycan (33–36). The addition of soluble heparin or heparan sulfate at concentrations ranging from 0.1 ng/ml to 10 μg/ml into the culture medium did not reconstitute the mitogenic response in chlorate-treated MCF-10A cells exposed to AR.[2] This strongly suggests that the HS chain(s) that are essential to AR-triggered activation of the EGFR need to be tethered to the cell surface.

Unlike EGF and transforming growth factor-α (TGF-α), but similar to HB-EGF, AR contains a very basic NH_2-terminal extension, relative to its EGF-like domain (1). The finding that a synthetic peptide corresponding to this region of AR (residues 26–44) binds to heparin-agarose and can compete for the binding of AR to heparin-agarose strongly suggests that this region of AR is at least partially involved in the interaction with HS which is required for AR-induced mitogenic signaling by the EGFR. Consistent with this observation is the fact that a homologous region in HB-EGF appears to be directly involved in the interaction between HB-EGF and heparin (50) and HS appears to be necessary for HB-EGF stimulation of smooth muscle migration (39). Therefore, it is plausible that mechanistically, HB-EGF may function in a manner similar to AR. Whether HS GAG is required for all the biological responses elicited by AR and HB-EGF remains to be seen.

Also contained within residues 26–44 of AR are two putative nuclear localization signals (1) and indeed, immunoreactive AR has been detected in the nucleus of cells *in vitro* and *in vivo* (6, 7, 9–13). The addition of exogenous ^{125}I-AR to several human carcinoma cell lines results in a preferential association of radiolabeled AR with nuclei, relative to radiolabeled EGF (51). Further, expression of SDGF (rat AR) lacking the secretory signal peptide results in nuclear accumulation of SDGF (52). It is possible that this very basic NH_2-terminal region of AR performs distinct functions depending upon whether AR is intracellular or extracellular. Lastly, the finding that an accessory molecule containing HS is needed for efficient AR action provides a mechanism by which AR may act in a more specific manner, relative to EGF and TGF-α. Since the EGFR is expressed on numerous different cell types *in vivo*, AR action, in contrast to that of EGF and TGF-α, may be specifically targeted to cells which co-express EGFR and the proper HS proteoglycan(s).

Acknowledgments—We thank Drs. William Gullick and Sally Prigent for generously supplying us with the AR^{68-84} peptide and Dr. Pier Paolo Di Fiore for generously providing us with the E7 antiserum.

[2] G. R. Johnson, unpublished observations.

REFERENCES

1. Shoyab, M., Plowman, G. D., McDonald, V. L., Bradley, J. G., and Todaro, G J. (1989) *Science* **243**, 1074–1076
2. Cook, P. W., Mattox, P. A., Keeble, W. W., Pittelkow, M. R., Plowman, G. D., Shoyab, M., Adelman, J. P., and Shipley, G. D. (1991) *Mol. Cell. Biol.* **11**, 2547–2557
3. Johnson, G. R., Prigent, S. A., Gullick, W. J., and Stromberg, K. (1993) *J. Biol. Chem.* **268**, 18835–18843
4. Johnson, G. R., Kannan, B., Shoyab, M, and Stromberg, K. (1993) *J. Biol. Chem.* **268**, 2924–2931
5. Shoyab, M., McDonald, V. L., Bradley, J. G., and Todaro, G. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6528–6532
6. Johnson, G. R., Saeki, T., Auersperg, N., Gordon, A. W., Shoyab, M., Salomon, D. S., and Stromberg, K. (1991) *Biochem. Biophys. Res. Commun.* **180**, 481–488
7. Johnson, G. R., Saeki, T., Gordon, A. W., Shoyab, M., Salomon, D. S., and Stromberg, K. (1992) *J. Cell Biol.* **118**, 741–751
8. Plowman, G. D., Green, J. M., McDonald, V. L., Neubauer, M. G., Disteche, C. M., Todaro, G. J., and Shoyab, M. (1990) *Mol. Cell. Biol.* **10**, 1969–1981
9. Saeki, T., Stromberg, K., Qi, C.-F., Gullick, W. J., Tahara, E., Normanno, N., Ciardiello, F., Kenney, N., Johnson, G. R., and Salomon, D. S. (1992) *Cancer Res.* **52**, 3467–3473
10. Kitadai, Y., Yasui, W., Yokozaki, H., Kuniyasu, H., Ayhan, A., Haruma, K., Kajiyama, G., Johnson, G. R., and Tahara, E. (1993) *Jpn. J. Cancer Res.* **84**, 879–884
11. Saeki, T., Salomon, D. S., Normanno, N., Johnson, G. R., Gullick, W. J., Mandai, K., Moriwaki, S., Takashima, S., Kuniyasu, M., Tahara, E., Kawami, H., Nishiyama, M., and Toge, T. (1994) *Int. J. Oncol.* **5**, 215–223
12. Qi, C.-F., Liscia, D. S., Normanno, N., Merlo, G., Johnson, G. R., Gullick, W. J., Ciardiello, F., Saeki, T., Brandt, R., Kim, N., Kenney, N., and Salomon, D. S. (1994) *Br. J. Cancer* **69**, 903–910
13. Ebert, M., Yokoyama, M., Kobrin, M. S., Friess, H., Lopez, M. E., Buchler, M. W., Johnson, G. R., and Korc, M. (1994) *Cancer Res.* **54**, 3959–3962
14. Ciardiello, F., Kim, N., Saeki, T., Dono, R., Persico, M. G., Plowman, G. D., Garriques, J., Radke, S., Todaro, G. J., and Salomon, D. S. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 7792–7796
15. Cook, P. W., Pittelkow, M. R., Keeble, W. W., Graves-Deal, R., Coffey, R. J., Jr., and Shipley, G. D. (1992) *Cancer Res.* **52**, 3224–3227
16. LeJeune, S., Leek, R., Horak, E., Plowman, G., Greenall, M., and Harris, A. L. (1993) *Cancer Res.* **53**, 3597–3602
17. Kimura, H., Fischer, W. H., and Schubert, D. (1990) *Nature* **348**, 257–260
18. Higashiyama, S., Abraham, J. A., Miller, J., Fiddles, J. C., and Klagsbrun, M. (1991) *Science* **251**, 936–939
19. Shing, Y., Christofori, G., Hanahan, D., Ono, Y., Sasada, R., Igarashi, K., and Folkman, J. (1993) *Science* **259**, 1604–1607
20. Holmes, W. E., Sliwkowski, M. X., Akita, R. W., Henzel, W. J., Lee, J., Park, J. W., Yansura, D., Abadi, N., Raab, H., Lewis, G. D., Shepard, H. M., Kuang, W.-J., Wood, W. I., Goeddel, D. V., and Vandlen, R. L. (1992) *Science* **256**, 1205–1211
21. Wen, D., Peles, E., Cupples, R., Suggs, S. V., Bacus, S. S., Luo, Y., Trail, G., Hu, S., Silbiger, S. M., Levy, R. B., Koski, R. A., Lu, H. S., and Yarden, Y. (1992) *Cell* **69**, 559–572
22. Peles, E., Ben-Levy, R., Tzahar, E., Liu, N., Wen, D., and Yarden, Y. (1993) *EMBO J.* **12**, 961–971
23. Carraway, K. L., III, Sliwkowski, M. X., Akita, R., Platko, J. V., Guy, P. M., Nuijens, A., Diamonti, A. J., Vandlen, R. L., Cantley, L. C., and Cerione, R. A. (1994) *J. Biol. Chem.* **269**, 14303–14306
24. Sliwkowski, M. X., Schaefer, G., Akita, R. W., Lofgren, J. A., Fitzpatrick, V. D., Nuijens, A., Fendly, B. M., Cerione, R. A., Vandlen, R. L., and Carraway, K. L., III (1994) *J. Biol. Chem.* **269**, 14661–14665
25. Plowman, G. D., Green, J. M., Culouscou, J.-M., Carlton, G. W., Rothwell, V. M., and Buckley, S. (1993) *Nature* **366**, 473–475
26. Ruoslahti, E. (1989) *J. Biol. Chem.* **264**, 13369–13372
27. Ruoslahti, E., and Yamaguchi, Y. (1991) *Cell* **64**, 867–869
28. Jackson, R. L., Busch, S. J., and Cardin, A. D. (1991) *Physiol. Rev.* **71**, 481–539
29. Esko, J., Rostand, K. S., and Weinke, J. L. (1988) *Science* **241**, 1092–1096
30. Bernfield, M., Kokenyesi, R., Kato, M., Hinkes, M. T., Spring, J., Gallo, R. L., and Lose, E. J. (1992) *Annu. Rev. Cell Biol.* **8**, 365–393
31. Yanagishita, M., and Hascall, V. C. (1992) *J. Biol. Chem.* **267**, 9451–9454
32. David, G. (1993) *FASEB J.* **7**, 1023–1030
33. Rapraeger, A. C., Krufka, A., and Olwin, B. B. (1991) *Science* **252**, 1705–1708
34. Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., and Ornitz, D. M. (1991) *Cell* **64**, 841–848
35. Klagsbrun, M., and Baird, A. (1991) *Cell* **67**, 229–231
36. Kan, M., Wang, F., Xu, J., Crabb, J. W., Hou, J., McKeehan, W. L. (1993) *Science* **259**, 1918–1921
37. Sakaguchi, K., Yanagishita, M., Takeuchi, Y., and Aurbach, G. D. (1991) *J. Biol. Chem.* **266**, 7270–7278
38. Gitay-Goren, H., Soker, S., Vlodavsky, I., and Neufeld, G. (1992) *J. Biol. Chem.* **267**, 6093–6098
39. Higashiyama, S., Abraham, J. A., and Klagsbrun, M. (1993) *J. Cell Biol.* **122**, 933–940
40. Di Fiore, P. P., Segatto, O., Lonardo, F., Faziola, F., Pierce, J. H., and Aaronson, S. A. (1990) *Mol. Cell. Biol.* **10**, 2749–2756
41. Farley, J. R., Nakayama, G., Cryns, D., and Segel, I. H. (1978) *Arch. Biochem. Biophys.* **185**, 376–390
42. Ullrich, A., and Schlessinger, J. (1990) *Cell* **61**, 203–212
43. Oike, Y., Kimata, K., Shinomura, T., Nakazawa, K., and Suzuki, S. (1980) *Biochem. J.* **191**, 193–207
44. Ototani, N., Kikuchi, M., and Yoshizawa, Z. (1981) *Carbohydr. Res.* **88**, 291–303

45. Li, S., Plowman, G. D., Buckley, S. D., and Shipley, G. D. (1992) *J. Cell. Physiol.* **153**, 103–111
46. Kenney, N., Johnson, G., Selvam, M. P., Kim, N., Qi, C.-F., Saeki, T., Brandt, R., Wallace-Jones, B., Ciardiello, F., Shoyab, M., Plowman, G., Day, A., Salomon, D. S., and Normanno, N. (1993) *Mol. Cell Differ.* **1**, 163–184
47. Normanno, N., Selvam, M. P., Qi, C.-F., Saeki, T., Johnson, G., Kim, N., Ciardiello, F., Shoyab, M., Plowman, G., Brandt, R., Todaro, G., and Salomon, D. S. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 2790–2794
48. Carey, D. J., and Stahl, R. C. (1990) *J. Cell Biol.* **111**, 2053–2062
49. Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Downward, J., Mayes, E. L. V., Whittle, N., Waterfield, M. D., and Seeburg, P. H. (1984) *Nature* **309**, 418–425
50. Thompson, S. A., Higashiyama, S., Wood, K., Pollitt, N. S., Damm, D., McEnroe, G., Garrick, B., Ashton, N., Lau, K., Hancock, N., Klagsbrun, M., and Abraham, J. A. (1994) *J. Biol. Chem.* **269**, 2541–2549
51. Modrell, B., McDonald, V. L., and Shoyab, M. (1992) *Growth Factors* **7**, 305–314
52. Kimura, H. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 2165–2169

The Journal of Biological Chemistry

MOLECULAR AND CELLULAR BIOLOGY, May 1990, p. 1969–1981
0270-7306/90/051969-13$02.00/0

Vol. 10, No. 5

The Amphiregulin Gene Encodes a Novel Epidermal Growth Factor-Related Protein with Tumor-Inhibitory Activity

GREGORY D. PLOWMAN,[1]* JANELL M. GREEN,[1] VICKI L. McDONALD,[1] MICHAEL G. NEUBAUER,[1] CHRISTINE M. DISTECHE,[2] GEORGE J. TODARO,[1] AND MOHAMMED SHOYAB[1]

Oncogen, 3005 First Avenue, Seattle, Washington 98121,[1] and Departments of Medicine (Medical Genetics) and Pathology, University of Washington, Seattle, Washington 98195[2]

Received 2 November 1989/Accepted 3 January 1990

We have isolated the gene for a novel growth regulator, amphiregulin (AR), that is evolutionarily related to epidermal growth factor (EGF) and transforming growth factor α (TGF-α). AR is a bifunctional growth modulator: it interacts with the EGF/TGF-α receptor to promote the growth of normal epithelial cells and inhibits the growth of certain aggressive carcinoma cell lines. The 84-amino-acid mature protein is embedded within a 252-amino-acid transmembrane precursor, an organization similar to that of the TGF-α precursor. Human placenta and ovaries were found to express significant amounts of the 1.4-kilobase AR transcript, implicating AR in the regulation of normal cell growth. In addition, the AR gene was localized to chromosomal region 4q13-4q21, a common breakpoint for acute lymphoblastic leukemia.

Cell growth and differentiation are regulated in part by the specific interaction of secreted growth factors and their membrane-bound receptors. Receptor-ligand interaction results in activation of intracellular signals leading to specific cellular responses. Epidermal growth factor (EGF), platelet-derived growth factor, insulin, insulinlike growth factor 1, colony-stimulating factor 1, and fibroblast growth factor all transmit their growth-modulating signals by binding to and activating receptors with intrinsic tyrosine kinase activity (reviewed in reference 30). Characterization of the physiologic and chemical effects that result from the binding of EGF and transforming growth factor α (TGF-α) to the EGF receptor has served as a useful model for understanding receptor-ligand interactions, signal transduction, and the regulation of cell growth and oncogenesis (reviewed in reference 76).

We have recently reported the purification and sequence analysis of a glycoprotein isolated from the conditioned media of 12-*O*-tetradecanoylphorbol-13-acetate (TPA)-treated MCF-7 cells (65, 66). The protein was termed amphiregulin (AR) to reflect its bifunctional activities: it inhibits the growth of many human tumor cells, and it stimulates the proliferation of normal fibroblasts and keratinocytes. The secreted protein exists as a monomer of either 78 or 84 amino acids (aa), with the shorter form lacking the six N-terminal residues of the larger molecule. Sequence analysis reveals that AR has a region with striking homology to EGF (38%) and TGF-α (32%), yet it also has an N-terminal extension of 43 aa composed primarily of very basic, hydrophilic residues (Lys, Arg, and Asn). In addition, AR has functional homology with this class of growth factors; it partially competes for binding of EGF to the EGF receptor and can supplant the need for EGF or TGF-α to maintain keratinocytes in culture. AR differs from EGF and TGF-α in that it fails to promote anchorage-independent growth of normal rat kidney (NRK) fibroblasts in the presence of TGF-β and inhibits the growth of certain tumor cells that proliferate in response to EGF or TGF-α.

In this report, we describe the isolation and characterization of cDNA and genomic clones for human AR, the transcriptional profile of the AR gene, and its chromosomal localization. Like EGF and TGF-α, AR is synthesized as a transmembrane precursor, with the secreted protein being released by proteolytic cleavage. AR is found in many of the same tissues (ovary, testis, and breast) and tumor types (squamous carcinomas and mammary adenocarcinomas) as TGF-α, but analysis of individual tumor cell lines reveals an inverse correlation between expression of TGF-α and AR. In addition, AR transcription is stimulated in several human breast cancer cell lines after treatment with TPA. These findings suggest that AR has novel growth-regulatory activities on both normal and neoplastic cells.

MATERIALS AND METHODS

Cell culture. All cells were obtained from the American Type Culture Collection. MCF-7 cells were maintained in 50% Iscove's modified Dulbecco medium–50% Dulbecco modified Eagle medium containing 10% heat-inactivated fetal bovine serum and 0.6 μg of insulin per ml. All other cell lines were grown in Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum.

cDNA cloning. Total cellular RNA was isolated from MCF-7 cells after treatment with 100 ng of TPA per ml for 24, 40, and 72 h by the guanidinium method. Poly(A)$^+$ RNA was isolated from pooled aliquots of these samples. First-strand cDNA synthesis was performed on 5 μg of poly(A)$^+$ RNA primed with oligo(dT) essentially as described by Gubler and Hoffman (29). Second-strand synthesis was performed with 4U of RNase H and 115 U of DNA polymerase I. T4 DNA polymerase (10 μg) was used for removal of 3′ overhangs, creating blunt ends. Double-stranded cDNA was sized over a Sephadex G-50 column to select for cDNAs longer than 500 base pairs (bp), and then 150 ng of cDNA was dG tailed with terminal deoxynucleotidyl transferase. dG-tailed cDNA was ligated into the *Eco*RI site of λgt10 by using the BR1 adapters (AATTCCCCCCCCCCCC) as described by Rose et al. (59). Duplicate nitrocellulose lifts were taken on 2.5×10^5 recombinants and filters were probed with best-guess (ARK31 and ARK41) and degenerate (ARD41 and ARD58) oligonucleotides derived from the human amphiregulin protein sequence (66). The oligonucleotide probes (including their degeneracy or length and corresponding

* Corresponding author.

amino acid residues) which were used for screening the λgt10 library were as follows: ARD41, 5′-TAYTCYTGYTG RCAYTTRCA-3′ (64-fold, CKCQQEY); ARD58, 5′-TGYT CAATRTAYTTRCAYTC-3′ (32-fold, ECKYIEH); ARK41, 5′-TCGCCACACCGCTCGCCAAAGTACTCCTGCTGGCA CTTGCAGGTCACAGCCTCCAGATGCTCAATGT-3′ (67-mer, YIEHLEAVTCKCQQEYFGERCGE); ARK31, 5′-TT GCACTCGCCATGGATGAGAAGTTCTGGAACTCAGCA TTGCATGGGTTCTT-3′ (53-mer, KNPCNAEFQNFCIHG ECK). The probes all correspond to the antisense mRNA strand, and degenerate resides are R = A or G, Y = C or T. Oligonucleotides were labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase (3 × 10^8 cpm/μg). The λgt10 cDNA library was first screened with a mixture of degenerate probes (ARD41 plus ARD58) and later reprobed with a mixture of nondegenerate probes (ARK31 plus ARK41). Hybridization was performed in oligonucleotide hybridization mix (6× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 5× Denhardt solution, 0.15% sodium PP_i, 0.1 mg of salmon sperm DNA per ml, 0.1 mg of tRNA per ml) at 37°C overnight. Washes were done in 6× SSC at 37°C. Under these conditions ARK31 failed to show a reproducible signal.

DNA sequence analysis. The 1.25-kilobase-pair (kb) cDNA insert of λAR1 was sequenced on both strands by the dideoxy-chain termination method (61) with specific oligonucleotide primers. Additional cDNAs were used to confirm the sequence of the coding region. Genomic DNA spanning the entire cDNA sequence, intron-exon junctions, and the 5′ regulatory region was also sequenced. No differences were detected between the genomic and cDNA sequences.

Northern (RNA) and Southern blot analyses. RNA was isolated from subconfluent cells grown in T150 tissue culture flasks or from fresh-frozen tissue samples. Total RNA (10 or 20 μg) was fractionated on 1.0% agarose–formaldehyde gels, transferred to nylon membranes (Amersham Hybond-N), and UV cross-linked (1,200 μJ on Stratalinker; Stratagene). Probes were prepared by random-prime ^{32}P-labeling (specific activity, 5 × 10^8 to 25 × 10^8 cpm/μg) of a 480-bp cDNA fragment (ARBP1) spanning the entire coding region of mature AR or a 170-bp cDNA fragment (AR170) encoding the transmembrane and cytoplasmic domains of AR (23). Hybridizations were performed in 5× SSPE (1× SSPE is 0.18 M NaCl, 10 mM sodium phosphate [pH 7.7], plus 0.1 mM EDTA)–5× Denhardt solution–0.5% sodium dodecyl sulfate (SDS)–20 μg of denatured salmon sperm DNA per ml at 42°C for 16 h with 2 × 10^6 cpm of the [^{32}P]ARBP1 per ml. Blots were washed several times in 2× SSC–0.1% SDS at 65°C and then in 1× SSC–0.1% SDS at 65°C and exposed on Kodak X-OMAT with two Du Pont Cronex Lightning Plus intensifying screens at −70°C.

Genomic DNAs were isolated from subconfluent cells in T150 tissue culture flasks. DNA (20 μg) was digested, analyzed on 0.8% agarose gels (SeaKem GTG), and blotted onto nylon membranes (Hybond-N). Filters were hybridized overnight at 42°C in Southern hybridization buffer (6× SSC, 5× Denhardt solution, 0.5% SDS, 20 μg of denatured salmon sperm DNA per ml) containing 2 × 10^6 cpm of ^{32}P-labeled AR-specific fragment per ml. Filters were washed extensively in 1× SSC–0.1% SDS at 65°C and autoradiographed overnight at −70°C.

Genomic cloning. MCF-7 DNA was digested with *Hin*dIII and electrophoresed on 0.8% low-gel-temperature agarose (Bio-Rad Laboratories). DNA was extracted from the agarose fractions spanning the 12- and 6.4-kb range and inserted into the *Hin*dIII site of the λL47.1 vector (45). Nitrocellulose plaque lifts were hybridized overnight at 42°C in Southern hybridization buffer containing 10% dextran sulfate. Filters were washed in 2× SSC–0.1% SDS at 65°C and then in 0.5× SSC–0.1% SDS at 65°C.

Primer extension analysis. Synthetic oligonucleotides AR(CP) and AR(AP) complementary to the AR 5′ cDNA sequence were ^{32}P end labeled with T4 polynucleotide kinase to a specific activity of 2 × 10^8 to 5 × 10^8 cpm/μg. Labeled oligonucleotide (10^6 cpm) was used to prime first-strand cDNA synthesis on 50 μg of MCF-7 RNA. The products were treated with RNase A, extracted with phenol and chloroform, ethanol precipitated, and analyzed by electrophoresis on standard 8% polyacrylamide–7 M urea sequencing gels. The sequences (and positions as numbered in Fig. 2) of the AR-specific oligonucleotides used for primer extension analysis are as follows: AR(AP), 5′-GCGGCG CCTCGGGCTGTCCCG-3′ (−151 to −171); AR(CP), 5′-CC GCTCTCGAAGGCTTGGGGAG-3′ (−114 to −135).

Plasmid constructions. For the promoter assays, plasmid pARS5, containing the 726-bp *Eco*RI-*Sst*I fragment from the 5′ end of the AR gene, was constructed (see Fig. 2). pARS5 was digested with *Sma*I and *Sph*I, blunted with S1 nuclease, and religated to generate pARS5Sm, which has the *Sma*I site (171 nucleotides upstream of the AR-initiating ATG) of the AR 5′ untranslated region (5′UTR) adjacent to a *Hin*dIII site. The 693-bp *Eco*RI-*Hin*dIII fragment was isolated from pARS5Sm and ligated into the *Hin*dIII site of the expression vector pSV0CAT by using a double-stranded oligonucleotide linker, ELNK3. The linker provides a *Hin*dIII-*Eco*RI adapter with internal *Kpn*I, *Aat*II, and *Sal*I sites to be used for generating exonuclease III deletions (33). The resulting expression construct, pXARE1CAT, contains 648 bp of AR 5′-flanking sequences, the cap site, and an additional 40 bp of exon 1 of the AR gene fused to the chloramphenicol acetyltransferase (CAT) gene. This plasmid was then linearized with *Kpn*I-*Sal*I, and constructs containing sequentially smaller amounts the AR 5′ region were generated by using the exonuclease III digestion method (33). The inserted DNA and flanking regions were verified by sequence analysis.

CAT assays. MCF-7 cells (2 × 10^6) were plated in 10 ml of medium in a 100-mm dish 12 to 16 h before transfection. The cells were transfected with 20 μg of calcium phosphate-precipitated supercoiled plasmid DNA, and after 4 h were subjected to a 25% glycerol shock for 90 s. They were then fed 20 ml of fresh medium containing 0 or 100 ng of TPA per ml. At 40 h after transfection, the cells were washed, collected, and lysed by sonication in 100 μl of 0.25 M Tris hydrochloride (pH 7.8). CAT activity was assayed essentially as described previously (27). Cell extract (3 to 50 μl) was added to 2.5 mCi of [^{14}C]chloramphenicol (Du Pont, NEN Research Products) in a 150-μl reaction volume containing 0.5 M Tris (pH 7.8) and 0.5 mM acetyl coenzyme A. The reactions were incubated at 37°C for 2 h, extracted with 1 ml of ethyl acetate, and developed on silica gel thin-layer chromatography plates with $CHCl_3$–1-butanol (95:5). The thin-layer chromatography plates were dried and autoradiographed. The acetylated and unacetylated [^{14}C]chloramphenicol was quantified in a scintillation counter. CAT enzymatic activity was calculated as micrograms of chloramphenicol acetylated per hour per milligram of protein in the cell extract.

Chromosomal localization. Plasmid pAR9, containing the complete AR cDNA sequence except for 100 bp from the 3′UTR, was random-prime labeled by using ^{3}H-nucleotides to a specific activity of 3 × 10^7 cpm/μg (23). In situ

hybridization to metaphase chromosomes from lymphocytes of a normal male donor was performed with AR probe concentrations of 10 to 30 ng/μl (49). The slides were exposed for 3 to 4 weeks, and chromosomes were identified by Q-banding.

RESULTS

AR cDNA. Amphiregulin was initially identified and sequenced from the supernatant of MCF-7 cells following 2 to 3 days of treatment with TPA (66). This human breast carcinoma cell line was then used as a source for cloning the AR cDNA. Two λgt10 clones were plaque purified based on positive hybridization to oligonucleotide probes derived from the AR protein sequence. The clones (λAR1 and λAR2) each contained a single 1.3-kb *Eco*RI insert and were sequenced with a degenerate oligonucleotide (ARD58) to verify that they specifically encoded a protein whose sequence matched that of a major portion of the AR protein.

A 170-bp fragment (AE170) from the AR cDNA clone was used to probe a second cDNA library of 200,000 recombinants. Thirteen positive recombinants were identified. The inserts ranged from 300 bp to 1.3 kb; six were longer than 1 kb, and five contained a single *Sst*I site known to be within 100 bp from the 5′ end of the longest cDNA clone. Four of these larger inserts were subcloned for further restriction and sequence analysis. All clones had identical restriction maps based on *Bsm*I, *Eco*RV, *Pvu*II, *Sst*I, and *Sma*I digestions, except one, which had a 3′ truncation of 100 bp and was later found to originate from an A_5 track, presumably sufficient for priming with oligo(dT).

Exact oligonucleotide primers were used to sequence both strands of the 1,230-bp AR cDNA (Fig. 1A). An open reading frame of 965 bp begins at nucleotide 1. The first AUG, at position 210, does not conform with the optimal consensus for translational start sites (37) owing to lack of a purine at position −3, although the second methionine (position 378) does match this consensus. The true start site is thought to be at the first AUG, since it is followed by a predicted 19-aa stretch of predominantly hydrophobic residues typical of a signal peptide sequence (72).

The cDNA encodes a protein precursor of 252 aa with a 210-bp 5′UTR. The translational termination signal (TAA) at position 757 is followed by a 262-bp 3′UTR. Comparison of the cDNA sequence with the nondegenerate probes showed 75% (ARK41) and 77% (ARK31) overall homology. Neither probe had a consecutive match of more than 8 nucleotides; however, 50 of 67 aligned nucleotides were sufficient to produce a detectable signal with ARK41 under conditions of low stringency. The codon usage by human AR mRNA sequence differs considerably from the usage frequencies reported by Lathe (41) and explains why the degenerate probes ARD41 and ARD58 showed stronger hybridization than the longer, preferred codon usage probes.

Hydropathy analysis of the AR precursor sequence revealed two hydrophobic domains and one extended hydrophilic stretch (Fig. 1B). The hydrophilic region is of notable length and magnitude, scoring below −4.0 with the algorithm of Kyte and Doolittle (39). Six structural domains are predicted in the 252 residue AR precursor: a 19-aa signal sequence (aa 1 to 19); an 81-residue amino-terminal domain (aa 20 to 100) which is serine rich (17 of 81 aa); an 84-residue region encoding the mature AR (aa 101 to 184), the first 43 aa being the hydrophilic domain and the last 41 aa showing homology to EGF and TGF-α; a hydrophobic, 23-residue putative transmembrane domain (aa 199 to 221) flanked by basic residues; and a 31-residue carboxy-terminal cytoplasmic domain (aa 222 to 252). The unglycosylated AR precursor minus the signal peptide has a predicted molecular weight of 25,942.

The 78- and 84-aa forms of mature AR are synthesized as the middle portion of a 252-aa transmembrane precursor. The cDNA sequence confirms the AR peptide sequence except for aa 113, which was sequenced as Asp (D) by protein analysis and was translated as Asn (AAC = N) from the cDNA sequence. The coding region was verified from three additional cDNA clones, and five cDNAs were all found to have their 5′ end within 25 bp of the longest clone. The cleavage sites for the release of the 78- and 84-aa forms of this growth factor do not correspond to those of known proteases, although a basic amino acid is located two residues downstream of both the N- and C-termini. Cleavage occurs between Asp-Asp and Ser-Val or between Glu-Gln and Val-Val at the amino termini and between Gln-Lys and Ser-Met at the carboxy terminus.

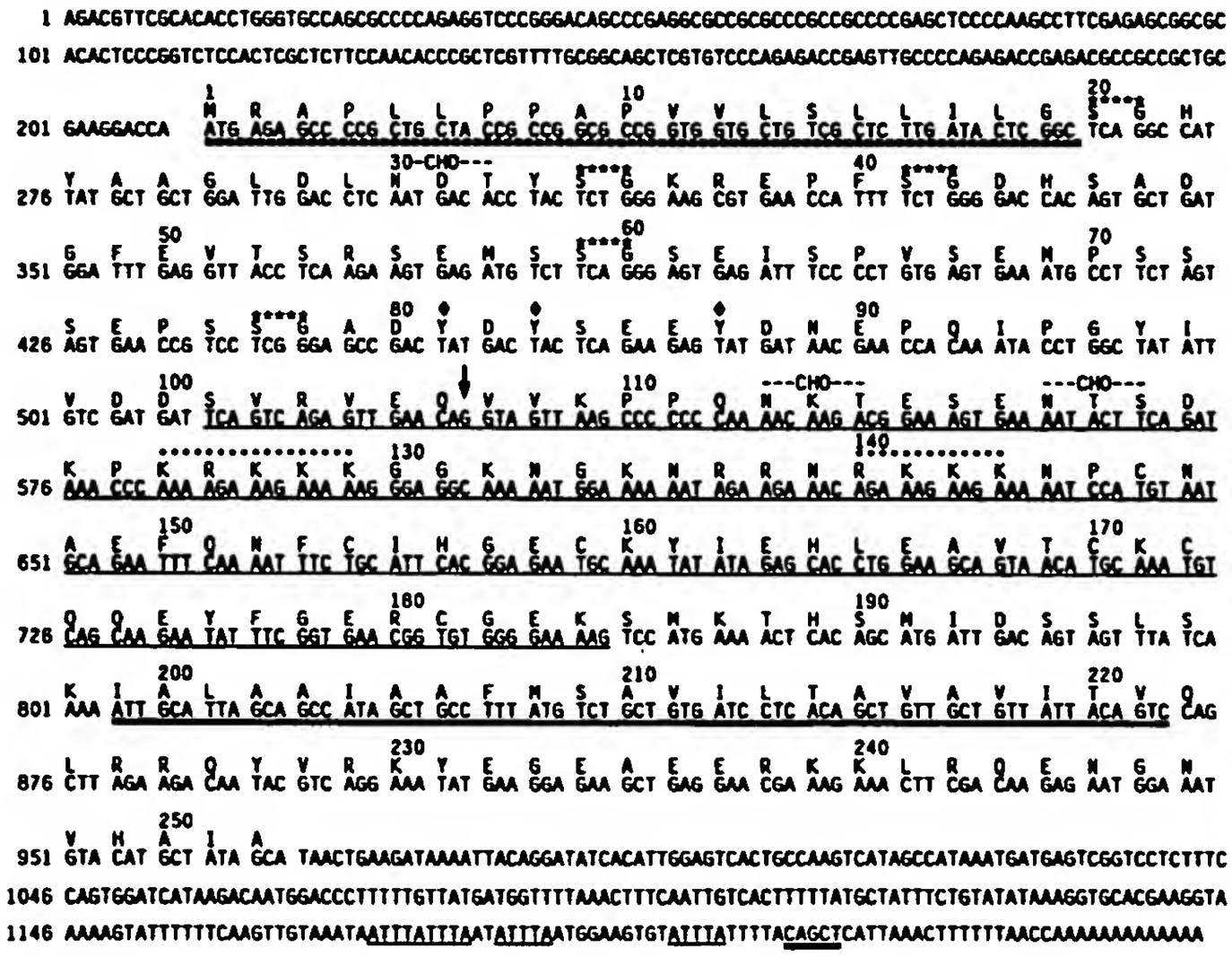
AR gene. Southern analysis of MCF-7 DNA digested with *Hin*dIII, *Eco*RI, or *Bam*HI showed single bands (12, 8, and >20 kb, respectively) when hybridized with a probe containing the 5′ portion of the AR cDNA (nucleotides 1 to 670). The absence of multiple bands suggests that AR is a single-copy gene. A further 3′ cDNA probe (nucleotides 681 to 850, spanning the transmembrane and cytoplasmic domains) hybridized to the same fragments and to an additional 6.4-kb *Hin*dIII fragment. These results indicate that the mature AR coding region is split between two *Hin*dIII fragments. Similar banding was seen on digests from human placenta, brain, melanoma (SK-MEL 28), choriocarcinoma (JEG-3), and epidermoid carcinoma (A-431) DNA, suggesting that there are no gross rearrangements or amplifications of the AR gene.

The two *Hin*dIII fragments were cloned from MCF-7 DNA, since these were likely to contain most, if not all, of the AR gene and flanking sequences. Of nine positive clones, two (λARH12 and λARH6) were selected for more detailed characterization. The 12- and 6.4-kb inserts were then subcloned and mapped for several restriction sites. The sequences of the exons and adjacent intron regions were determined with exact oligonucleotide primers combined with direct sequencing of smaller subclones. The genomic sequence confirmed the transcribed sequence that was determined from the cDNA clones (Fig. 2).

The human AR precursor is contained in six exons, spanning 10.2 kb of genomic DNA (Fig. 2 and 3). The exons vary in size from 112 to 270 bp and are interrupted by introns ranging from 1.25 to 2.1 kb. The intron-exon boundaries of the AR gene all conform with the canonical splice consensus sequence. The five introns of AR interrupt the coding sequence near the borders of the various protein domains (Fig. 3). Exon 1 encodes the 5′UTR and signal peptide; exon 2 encodes the N-terminal precursor; exon 3 encodes the very basic and hydrophilic N-terminal portion of AR as well as the first two loops of the EGF-like region; exon 4 comprises the third loop of the EGF-like motif and the transmembrane domain; exon 5 contains the cytoplasmic region; and exon 6 represents the 3′UTR. There is no intron separating the hydrophilic domain from the EGF-like motif, nor are any cryptic splice sites apparent between these domains. This hydrophilic region is purine rich (55 of 59 residues are A or G) and has no repetitive sequences. The process by which this unusual N-terminal extension became juxtaposed with the EGF-like region remains unclear.

Analysis of the AR 5′ regulatory region. The recombinant

A.



B.

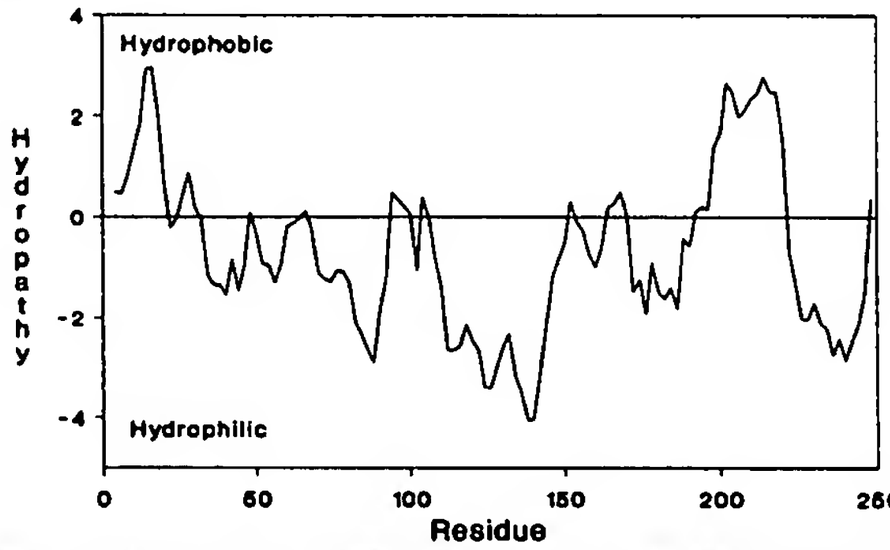


FIG. 1. Nucleotide sequence of the human AR cDNA and predicted amino acid sequence of the AR precursor. (A) Nucleotide numbering is on the left, and amino acid numbering is above every tenth residue. The 84 residues corresponding to the mature secreted AR protein are underlined, the predicted hydrophobic signal peptide is underscored with a dashed line, and the transmembrane sequence is marked with a double underline. An arrow marks the alternate N-terminal cleavage site for mature AR. Potential N-linked glycosylation sites are denoted by -CHO-, potential *N*-glycosaminoglycan attachment sites are identified by asterisks, consensus tyrosine sulfation sites are shown with a diamond, and basic residues conforming to a nuclear localization sequence are overscored with a dotted line. An exact AATAAA polyadenylation splice site is not present in the 3'UTR; however, a common polyadenylation site, CAGCT, is located 23 bp preceding the poly(A)$^+$ tract, indicated by the double underscore. Also underscored in the 3'UTR are the mRNA-destabilizing consensus motifs (ATTTA) characteristic of several cytokines and lymphokines (64). (B) Hydropathy profile of the predicted amino acid sequence of human AR precursor based on the algorithm of Kyte and Doolittle (39).

phage λARH12 contains 6.5 kb of genomic DNA 5′ to the cDNA clones. To better understand factors that govern the expression of AR, we localized the 5′ end of the AR gene by using primer extension analysis. Endogenous AR transcripts were characterized by using mRNA isolated from TPA-stimulated MCF-7 cells. Primer extension revealed three major transcriptional start sites between positions −211 and −209 (Fig. 4A), with the first located just 1 bp upstream of the longest cDNA clone (Fig. 2). A second oligonucleotide primer confirmed the extension products at positions −211 and −210.

To determine whether the AR 5′-flanking sequences con-

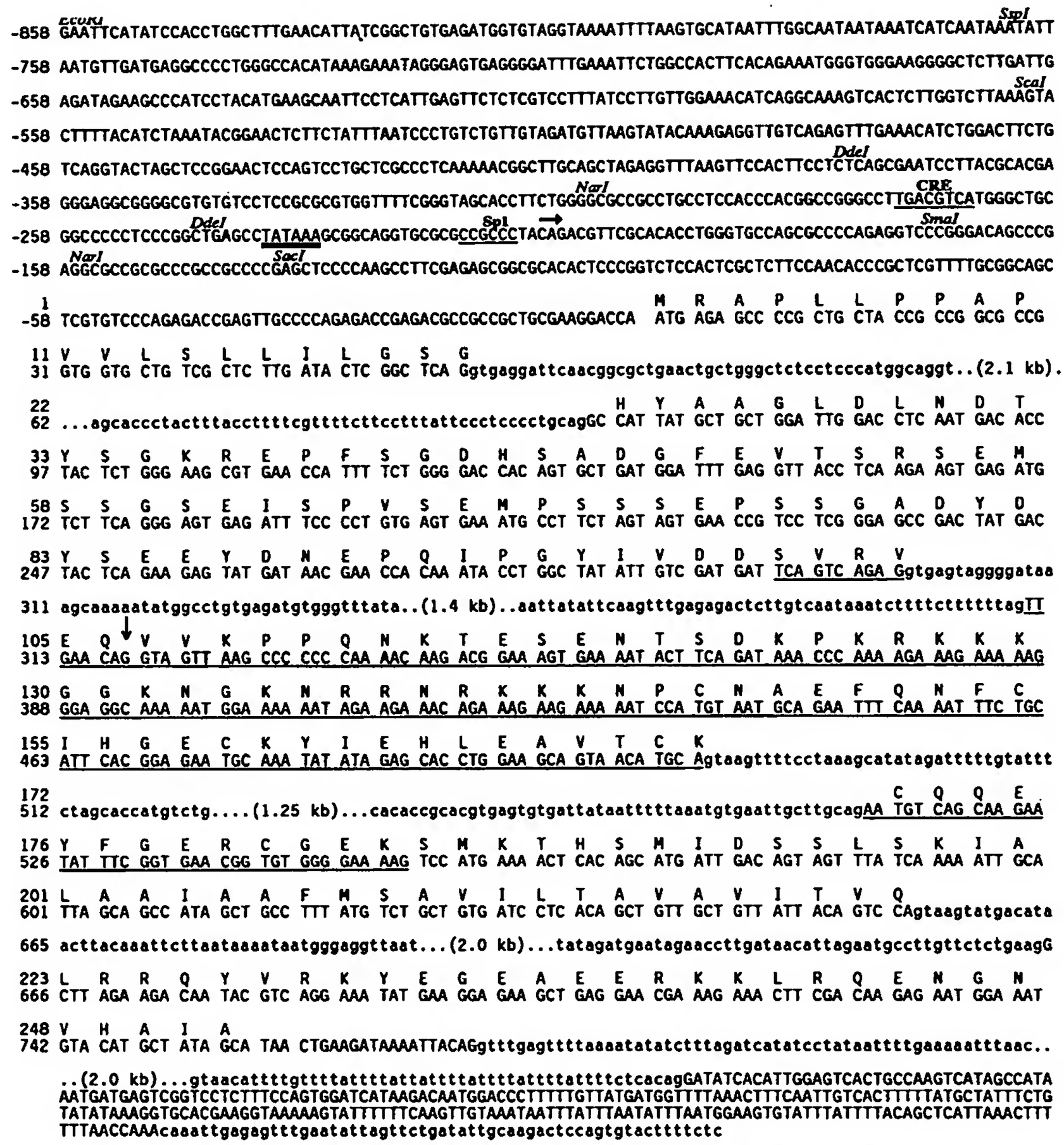


FIG. 2. Genomic and primary amino acid sequence of human AR. The sequence was determined from the two genomic *Hin*dIII fragments isolated from MCF-7 cells. The six exons are shown in capitals with introns and flanking sequences in lowercase letters. Nucleotide and protein sequences are numbered on the left, and the lengths of unsequenced intron regions are shown in parentheses. The mature AR coding region is underlined. The sequence of the AR promoter is also shown with numbering in reverse up to the *Eco*RI site at position −859. Selected restriction sites are shown as reference points. The furthest 5′ end of the mRNA is marked with a forward arrow, and an upstream TATA sequence is indicated by a double underline. An Sp1 consensus sequence and cAMP-responsive element (CRE) are underlined.

tain an active promoter, a series of vectors were constructed that contain the bacterial CAT gene under the control of various fragments from the AR 5′-flanking region (Fig. 4B). A transient-expression assay was used to assess promoter strength after transfection of the CAT vectors into MCF-7 cells (27). Plasmid pXARE1CAT, containing the 688-bp *Eco*RI-*Sma*I fragment from AR 5′-flanking sequences fused to CAT, showed detectable promoter activity that increased five- to eightfold in response to 40-h treatment with TPA. A similar plasmid, pXARE1dCAT, containing 167 bp of AR 5′-flanking sequences, showed activity comparable to pXARE1CAT and was also TPA inducible (Fig. 4B and C). Nuclear runoff experiments confirm three- to fivefold-increased transcription of AR in MCF-7 cells in response to TPA (data not shown). Taken together, these findings functionally identify a TPA-responsive promoter in the 5′-flanking region of AR gene.

The nucleotide sequence of an 859-bp genomic fragment derived from plasmid pARE6 is presented in Fig. 2. The 155-bp sequence immediately upstream of the 5′ end of the

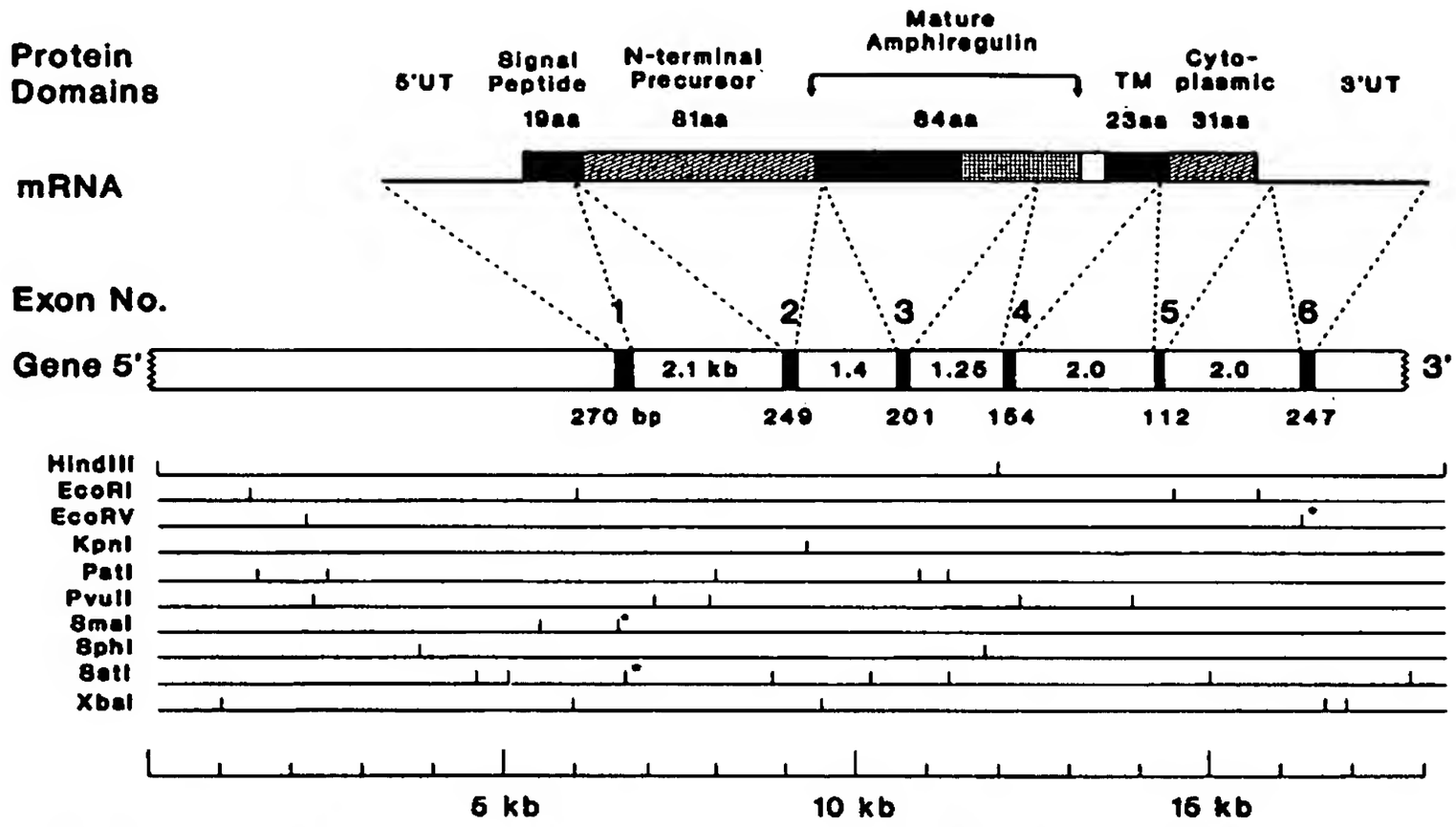


FIG. 3. Map of the human AR gene showing the exon organization and protein domains. The gene is drawn to scale in a 5'-to-3' orientation. The six exons are shown (▮), with the length in base pairs being listed directly under each exon. Intron lengths in kilobase pairs are indicated between each exon. Cleavage sites for 10 selected restriction enzymes are shown. Asterisks denote sites which are present in the cDNA. The corresponding position of each exon about the AR mRNA is shown above on a 15-fold-larger scale. Protein domains are represented by shaded boxes. The number of amino acid residues in each domain is indicated. The two dark filled boxes represent hydrophobic stretches that correspond to the signal peptide and transmembrane (TM) domains. Mature AR is represented by two boxes, the N-terminal hydrophilic domain and the C-terminal EGF-like motif.

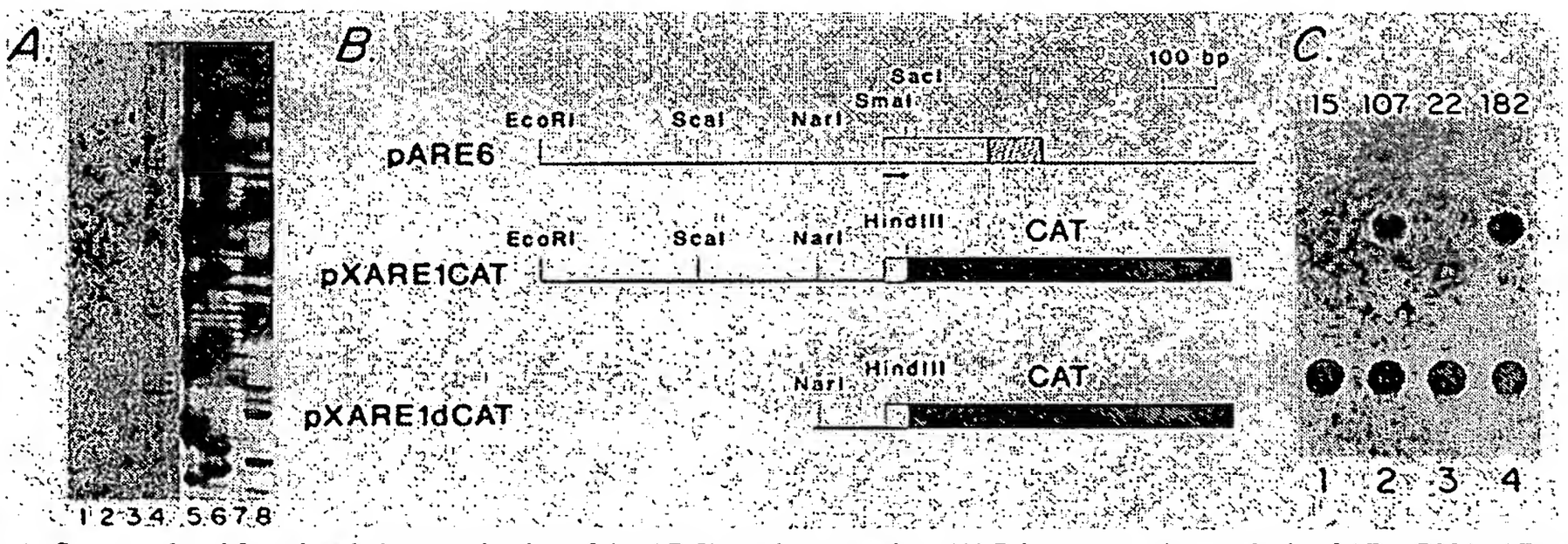


FIG. 4. Structural and functional characterization of the AR 5' regulatory region. (A) Primer extension analysis of AR mRNA. AR-specific primers were hybridized to 50 μg of total cytoplasmic RNA from MCF-7 cells (lane 1) or MCF-7 cells treated with TPA for 24 h (lanes 2 to 4). Lanes 1 and 2 contained primer AR(CP) at 5×10^5 cpm per reaction, whereas lanes 3 and 4 contained primer AR(AP) at 5×10^4 or 5×10^5 cpm per reaction, respectively (see Materials and Methods). cDNA was extended and subjected to electrophoresis on an 8% polyacrylamide–7 M urea gel. A dideoxy-sequencing ladder using primer AR(CP) on plasmid pARE6 is shown in lanes 5 to 8 (CTAG). Dots label the extended products, which were 96 and 97 nucleotides in lane 2 and 59, 60, and 61 nucleotides in lanes 3 and 4. Extensions from two separate primers concur on positions −211 and −210 as the AR transcriptional start sites. (B) Chimeric AR-CAT constructions. pARE6 is an AR genomic clone containing exon 1 and flanking sequences. Exon 1 is shown (▭), with an arrow indicating the direction of transcription, and the coding region is also shown (▨). A 689-bp *Eco*RI-*Sma*I fragment containing 41 bp of AR exon 1 and 5'-flanking sequences was inserted in front of the CAT gene (■) to generate pXARE1CAT. Exonuclease III (33) was used to delete 522 bp from the 5' end of this construct, retaining 167 bp of AR 5'-flanking DNA in plasmid pXARE1dCAT. (C) Induction of CAT enzyme activity in MCF-7 cells following transfection with pXARE1CAT (lanes 1 and 2) or pXARE1dCAT (lanes 3 and 4) and treatment with TPA for 40 h (lanes 2 and 4). An autoradiogram of a thin-layer chromatography plate is shown, with the acetylated products migrating above the unacetylated chloramphenicol. Above each lane is the amount of chloramphenicol (micrograms) acetylated per picogram of protein per hour. These results have been replicated four times.

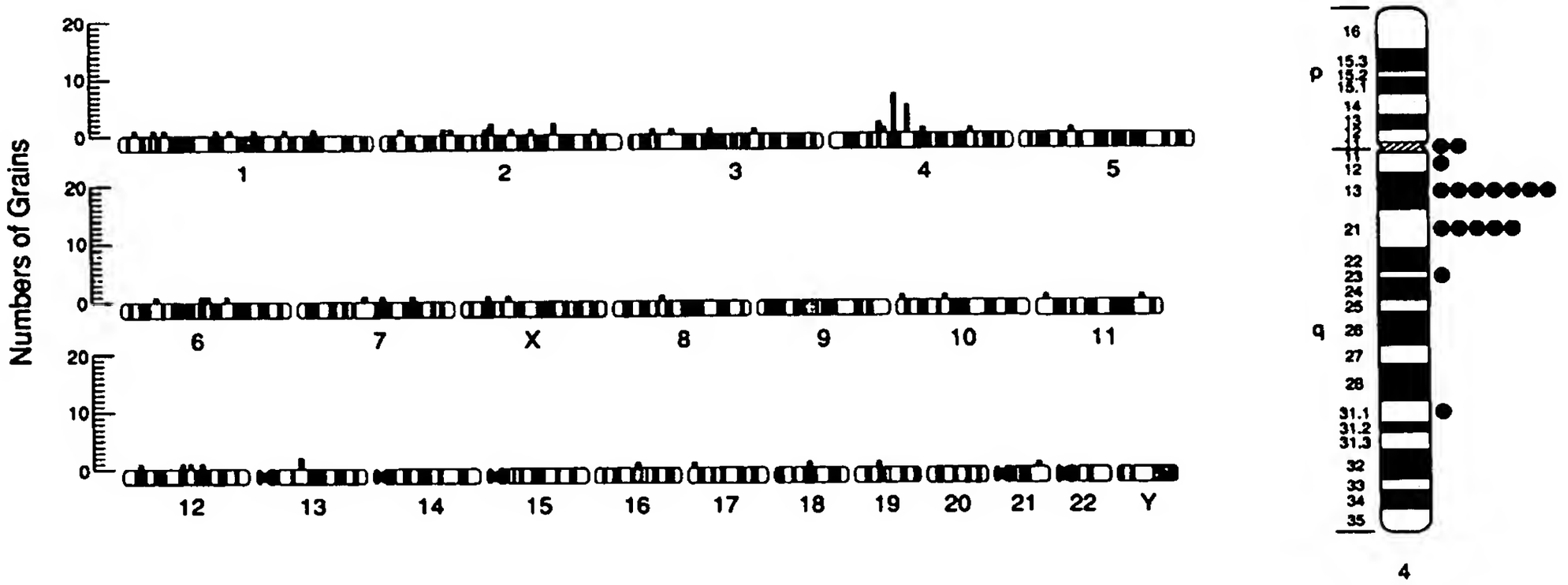


FIG. 5. The AR gene is located on human chromosome 4 band q13. (a) Distribution of 66 sites of hybridization on metaphase spreads of human chromosomes. (b) Regional localization of the AR gene on human chromosome 4 at 4q13-21. Shown is a histogram of 17 silver grains localized to chromosome 4. A nonrandom distribution of grains was observed, with a peak at bands 4q13-21.

mRNA has a high G+C content (73%) and a single Sp1-binding site. A consensus TATA sequence resides 28 bp upstream of the first mRNA start site, but no CAAT box is apparent. A potential cyclic AMP (cAMP)-responsive element (5'-TGACGTCA-3' [19]) is present beginning 64 bp upstream of the mRNA, and studies are under way to assess its significance. A consensus sequence often found in TPA-inducible genes [5'-TGA(G/C)TCA-3'] is not present in this region, suggesting that TPA may stimulate AR gene expression through a nonclassical pathway (10).

Chromosomal location of the AR gene. The AR gene was mapped on human metaphase chromosomes by in situ hybridization (49) with a cDNA clone as a ^{3}H-labeled probe. Forty-two metaphase cells were examined in this analysis. Of 66 sites of hybridization scored, 12 (18%) were located between bands q13 and q21 of the long arm of chromosome 4 (Fig. 5A). The largest number of grains was at band q13 (Fig. 5b). There was no significant hybridization to other chromosomes. Localization was confirmed by polymerase chain reaction on hamster × human somatic cell hybrid DNA containing only human chromosome 4. Oligonucleotide primers derived from AR exon 3 and the flanking intron generated a 220-bp polymerase chain reaction fragment only in human DNA and the somatic cell hybrid DNA containing chromosome 4; in contrast, the Chinese hamster ovary (CHO) DNA was negative (data not shown).

Cellular sources of AR transcripts. Northern analysis revealed a single 1.4-kb mRNA species for AR, which was detected in both total and poly(A)$^+$ RNA fractions from a variety of human tissues and tumor cell lines. The 1.4-kb transcript was most prevalent in RNA derived from human ovary and placenta and less abundant in pancreas, cardiac muscle, testis, colon, breast, lung, spleen, and kidney RNA (Fig. 6A). Little or no hybridization was seen in the adrenal, brain, duodenum, epidermis, liver, parathyroid, prostate, or thymus. AR-specific hybridization also was quantified by solution hybridization with ^{32}P-antisense RNA probes (35). This analysis revealed 3- to 10-fold-higher AR expression in the ovary and placenta than in the pancreas or spleen.

Several human cell lines were examined for expression of AR based on the presence of the 1.4-kb AR-homologous transcript and on solution hybridization. AR was expressed in six human breast carcinomas (MCF-7, MDA-MB-361, T-47D, BT-474, MDA-MB-231, and SKBR-3), but not in

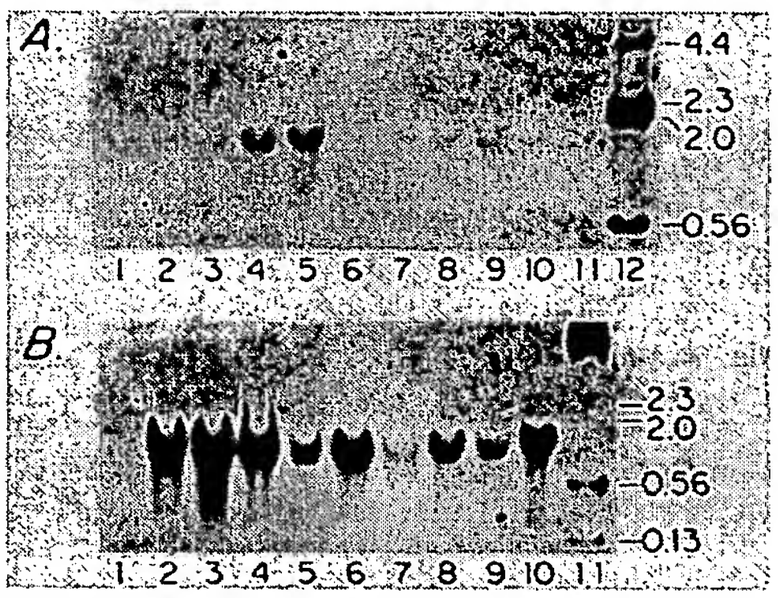


FIG. 6. Northern blot analysis of AR expression in normal human tissues and tumor cell lines. (A) Total RNA was isolated from the following human tissues and probed with an AR cDNA fragment. Lanes: 1, pancreas; 2, cardiac muscular; 3, testis; 4, placenta; 5, ovary; 6, epidermis; 7, duodenum; 8, colon; 9, breast; 10, cerebral cortex; 11, adrenal. Lane 12 is radiolabeled λ/*Hin*dIII markers with sizes shown in kilobases. All lanes contained 30 μg of total cytoplasmic RNA, except lanes 4, 5, and 9, on which 15 μg was loaded. The exposure time was 3 days at −70°C. (B) TPA induction of AR mRNA expression in tumor cell lines. Lanes 1 to 4 show a time course for TPA induction of AR in MCF-7 cells. RNA was isolated after 0 h (lane 1), 3 h (lane 2), 6 h (lane 3), or 24 h (lane 4) of exposure to 100 ng of TPA per ml. All other TPA treatments were for 24 h. Lane 5, MDA-MB-361 untreated; lane 6, MDA-MB-361 plus TPA; lane 7, BT-474 plus TPA; lane 8, T-47D plus TPA; lane 9, CaKi-1 plus TPA; lane 10, AsPC-1 plus TPA; lane 11, λ/*Hin*dIII markers. RNA was loaded on lanes 1 to 6 at 10 μg per lane and on lanes 7 to 10 at 20 μg per lane.

three others (MDA-MB-157, MDA-MB-468, and MDA-MB-453). AR was also detected in one of three kidney carcinomas (CaKi-1), two of four pancreatic carcinomas (AsPC-1 and Capan-1), and two of three ovarian tumors (CaOv-3 and SK-OV-3). Several embryonic (HEL-299, HEPM, JEG-3, HUF, and PA-1), melanoma (SK-MEL-28 and A-375), neural (CRL 7386 and SK-N-MC), and hematopoietic (CEM, U-937, and HSB-2) lines were negative. The highest levels of AR expression were seen in breast carcinoma cell lines that are estrogen receptor positive and contain low levels of EGF receptor (MCF-7, MDA-MB-361, T-47D, and BT-474) (Fig. 6B). TGF-α is also produced by many human tumor cells and has been detected in 70% of primary breast tumors (2); however, unlike AR, expression of TGF-α correlated with overexpression of the EGF receptor (17, 20). The four breast carcinoma cell lines that overexpressed AR (MCF-7, MDA-MB-361, T-47D, and BT-474) had low or undetectable levels of TGF-α (20, 52, 79; G. Plowman, unpublished observation). We conclude that AR is overexpressed in several mammary tumor cell lines, although its regulation is quite different from that of TGF-α.

The time course of AR induction by TPA was determined for MCF-7 and HBL-100 cells, the latter being a nonmalignant, estrogen receptor-negative, and progesterone receptor-negative breast cell line. RNA was isolated at 0, 3, 6, 24, and 48 h after TPA treatment and subjected to Northern and solution hybridization analysis. MCF-7 cells reproducibly showed an 11-fold increase in AR mRNA by 3 h with a maximum 18-fold increase by 24 h (Fig. 6B). In contrast, HBL-100 cells expressed detectable levels of AR only at the 1-h time point (data not shown). Of 18 human cell lines, 9 (5 of breast carcinoma origin) also showed increased AR expression in response to 24-h treatment with TPA. MDA-MB-361, a breast adenocarcinoma cell line with brain metastasis, showed the highest constitutive level of AR (21 pg of AR RNA per μg of total RNA), whereas MCF-7 cells showed the highest level after TPA induction (109 pg of AR RNA per μg of total RNA).

Because the AR 3′UTR contains A+U-rich sequence elements common to a number of genes which undergo rapid turnover (64), the stability of AR mRNA was analyzed in both the presence and absence of TPA. MCF-7 cells were treated with 5 to 40 μg of dactinomycin per ml, and cytoplasmic RNA was extracted at 0, 1, 2, 4, and 6 h. The decay of AR mRNA was monitored by Northern analysis and solution hybridization assays. A 6-h treatment with the lower dose of dactinomycin had no effect on AR mRNA levels. With higher concentrations of the drug, the AR half-life was shown to be 3.5 h, in both the presence and absence of TPA. The requirement for higher dactinomycin doses is probably due to delayed drug uptake by MCF-7 cells or to the inherent drug resistance common to many tumor cell lines (38). Despite the multiple ATTTA consensus sequences present in its 3′UTR, we conclude that the AR mRNA is quite stable in MCF-7 cells. Although similar motifs are reported to promote mRNA decay in normal hematopoietic cells, they may have a stabilizing effect in certain tumor lines (63). Multiple mechanisms are involved in the TPA stimulation of AR gene expression in MCF-7 cells; these include a two- to eightfold-increased transcription and the subsequent accumulation of a relatively stable mRNA. Additional studies are under way to ascertain whether similar mechanisms modulate AR gene expression in other cell types.

DISCUSSION

Structural comparison of AR and EGF-like growth modulators. A search of the amino acid (Protein Identification Resource, release 21) and nucleotide sequence (EMBL, release 19; GenBank, release 60) libraries was performed with the AR sequence. Only the region spanning the mature growth factor showed significant homology with any of the sequences in the data bases. EGF was most closely related to AR, with identical residues at 16 of the 37 (43%) positions in the region spanning the three disulfide loops (62). In the same region, TGF-α (47) and AR have 12 residues in common, TGF-α and EGF have 18, and EGF and vaccinia virus growth factor (7) have 17 (Fig. 7). The homology between these growth modulators includes the conserved spacing of the six cysteines involved in three disulfide bonds that define their secondary structure. Only four additional residues are completely conserved between AR and the other mammalian and viral EGF-like growth factors (Gly-13, Tyr-32, Gly-34, and Arg-36 [Fig. 7]; note that positions referred to in the text are relative to the first cysteine in mature EGF, as shown in Fig. 7A). Most proteins that bind the EGF receptor also contain His-11 and Gly-31; the exceptions are the viral EGF homologs from Shope fibroma virus and myxoma virus, which have Asn-11 and Glu-31 (9, 71), and AR, which has Glu-31. Residue 31 of the EGF-like proteins is located within a predicted type II β-turn, where such substitutions may be considered conservative (74). Two-dimensional nuclear magnetic resonance studies of EGF and TGF-α (11, 50) suggest that the highly conserved spacing of cysteines and glycines may define the basic structure of the protein backbone while other conserved, or conservatively changed, residues may form the functional receptor recognition site (residues 8, 11, 36, 38, and 42). Further support for their involvement in receptor binding is provided by analyses suggesting that these residues may all lie on the same face of the molecule (8).

Functional evidence that the conserved residues are necessary for biological activity has been obtained by characterization of derivatives of TGF-α and EGF. Recombinant proteins, synthetic peptides, site-specific chemical derivatives, and proteolytic degradation have all been useful for generating altered molecules (14, 31, 42). Some generalizations include the following: (i) the six cysteines (positions 1, 9, 15, 26, 28, and 37) and their disulfide loops (positions 1 to 15, 9 to 26, and 28 to 37) and Arg-36 are required for biological activity; (ii) N-terminal extensions have little effect on activity; (iii) an aromatic residue (Phe, Tyr) is required at position 8; and (iv) a nonconservative change of Tyr-32, Asp-41, or Leu-42 results in loss of activity and/or dramatic loss in receptor binding and autophosphorylation.

AR fits all but criterion (iv), since it truncates at position 40 and consequently lacks the final two conserved residues (Asp-41 and Leu-42). Despite this, AR competes for EGF receptor binding and substitutes for EGF in some mitogenic assays (66). These carboxy-terminal differences may be responsible for nonsaturable receptor-binding kinetics and the functional differences between AR and EGF, including an inability to synergize with TGF-β (66), marked differences from EGF and TGF-α in the inhibition of certain cell lines, and differences in cross-linking and phosphorylation assays (data not shown). The extended hydrophilic region at the N terminus of mature AR may also account for the disparities in receptor-binding and biological activity.

Structural domains that are superficially similar to EGF are found in diverse proteins, including several of the blood

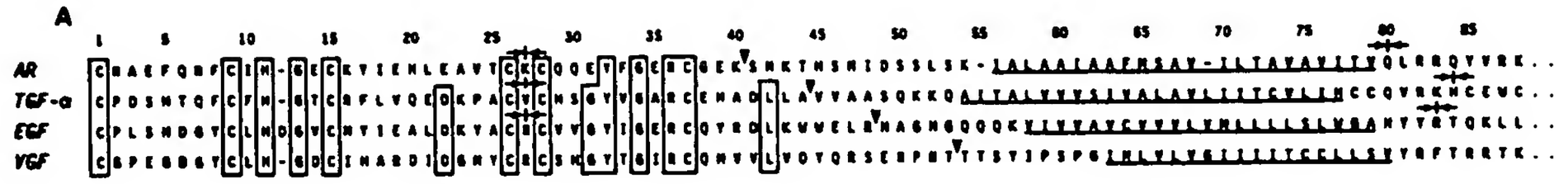


FIG. 7. Protein sequence homologies between the EGF-like growth modulators. (A) Amino acid alignment of the EGF-like motif and flanking transmembrane domain from three human proteins and one viral protein known to bind the EGF receptor. Alignment and numbering begins at the first cysteine of these motifs, and the most highly conserved residues are boxed. The putative transmembrane domains are underlined, and arrowheads mark the proteolytic cleavage sites where the mature growth modulators are released from their membrane-bound precursors. Exon-intron boundaries are displayed as facing arrows situated above the interrupted amino acids. The 3′ junction of human TGF-α is inferred from that of the rat gene. Vaccinia virus growth factor (VGF) contains no introns. (B) Schematic diagram of the predicted secondary structure of human AR, EGF, TGF-α, and vaccinia virus growth factor. Symbols: ●, residues which are completely conserved among all four proteins; ⊖, additional residues that EGF, TGF-α, or vaccinia virus growth factor have in common with AR; ◍, hydrophilic residues in the N-terminal portion of AR. An arrow marks the alternate cleavage site for the 78-aa form of mature AR.

coagulation factors; extracellular matrix proteins (laminin and tenascin); the invertebrate homeotic proteins including Notch, slit, and Delta from *Drosophila melanogaster* and *lin-12* and *glp-1* from *Caenorhabditis elegans*; low-density lipoprotein receptor and low-density lipoprotein receptor-related protein; and members of a family of cell adhesion proteins including the lymphocyte homing receptors and the endothelial leukocyte adhesion molecule ELAM-1 (reviewed in references 40, 66, and 77). Although structurally similar to EGF, none of these proteins maintain the precise spacing of all six cysteines, and, likewise, none have been shown to compete for binding to the EGF receptor.

AR precursor. Mature, secreted AR is synthesized as the middle portion of a 252-aa transmembrane precursor. The AR precursor has three potential N-glycosylation sites, one in the N-terminal domain (position 30) and two in the hydrophilic region of the mature protein (positions 113 and 119). The 78- and 84-aa forms of AR have predicted molecular weights of 9,173 and 9,772, respectively. N-linked glycosylation is known to contribute 10 to 12 kilodaltons to mature AR and is probably the result of carbohydrate addition at one or both of the sites in the hydrophilic domain.

The N-terminal region of the AR precursor contains 19 tightly clustered serine and threonine residues, 18 acidic residues, and 6 prolines. These attributes are common to O-glycosylation sites (48) and suggest that the AR precursor may contain complex carbohydrates, a modification observed in some larger forms of the TGF-α precursor (6, 25, 69). This domain also contains four Ser-Gly dipeptides that are potential sites for glycosaminoglycan attachment (26). Moreover, three potential tyrosine sulfation sites (Tyr-81, Tyr-83, and Tyr-87) are identified in the AR precursor based on the presence of a tyrosine residue surrounded by acidic amino acids and the absence of residues that could contrib-

ute to steric hindrance (1, 34). Residues 55 to 102 scored high in the PEST algorithm (58), which identifies protein regions rich in proline, glutamic acid, serine, and threonine residues. Proteins scoring high by this analysis typically undergo rapid degradation.

The hydrophilic domain of the mature AR protein is composed of many positively charged amino acids (16 of 43 residues are Lys or Arg), including two consecutive stretches of four or five basic residues (Fig. 1A). This region of AR is similar to the nuclear targeting signal of simian virus 40 large T antigen that contains a characteristic KKKRK sequence preceded by small amino acids (Gly, Ala, and Pro) thought to favor the formation of an α-helical structure (reviewed in reference 57). Mutation analysis has revealed four consecutive basic residues as the predominant feature of the simian virus 40 nuclear localization sequence. Other proteins that contain similar nuclear targeting sequences include histones, steroid hormone receptors, c-*abl*, MyoD1, and c-*myc*. Preliminary data show that AR binds single- and double-stranded DNA under conditions in which EGF does not bind (M. Shoyab and G. Plowman, unpublished data). Production of AR-specific antibodies (24) will be useful for localization of internalized AR. Possibly AR mediates some of its effects by being targeted to the nucleus and interacting with the controlling regions of other growth-regulatory genes.

Differential proteolytic cleavage of a transmembrane precursor is thought to generate the two forms of mature AR (78 and 84 aa). The presence of an intron between the predicted N-terminal cleavage sites suggested that intron sliding might account for this difference. MCF-7 cells produce 80% of their AR in the smaller (78-aa) form, whereas JEG-3, a choriocarcinoma cell line, produces most of its protein as the larger (84-aa) form (data not shown). To determine whether this difference was due to alterations at the DNA level, the 5′ junction of AR exon 3 was isolated from three human cell lines (JEG-3, CaKi-1, and MCF-7) by using polymerase chain reaction techniques. Direct sequence analysis revealed no cell-type-specific differences at the genomic level and supports a model of differential proteolytic processing for generation of the two forms of secreted AR.

Conservation of exon structure among AR, EGF, and TGF-α. The human AR gene is divided into six exons, spanning 10.2 kb (Fig. 2 and 3). Like AR, the TGF-α gene has six exons (16, 18), yet it spans nearly 10 times the length of the AR gene. The human EGF gene is also relatively large, with 24 exons covering 110 kb (3, 28). The exon organization of the EGF-like motifs of these three proteins is conserved. Exons 3 and 4 of AR encode the mature protein, with an intron disrupting the coding sequence between the second and third disulfide loops. Alignment of the amino acid sequences for AR, EGF, and TGF-α reveals an identically placed intron in all three proteins (Fig. 7A). Moreover, the adjacent exon of each contains the transmembrane domains of the precursor proteins. This configuration is associated with secretion of an active EGF receptor-binding protein and suggests that the integral membrane form may be necessary for efficient folding of the disulfide bonds. In addition, recent evidence suggests that these transmembrane precursors may be biologically active even in the absence of processing (5, 75). In contrast to the pattern for the three growth regulators, the cysteine-rich domain is contained on a single exon in all other homologous mammalian proteins for which the exon structure has been determined, including the eight EGF-like repeats in the human EGF precursor (3), and three repeats in the low-density lipoprotein receptor (68), and the single repeat in the tissue-type plasminogen activator, urokinase, and each of the human coagulation factors IX, X, XII, and protein C (reviewed in references 12 and 44).

Differences in the structures of these EGF-like repeats suggest that they may have distinct origins. One group contains the motif bounded by introns, whereas the other group, of which AR, EGF, and TGF-α are members, has the motif interrupted by an intron at a precise location. The sequence similarity between the two groups could be the result of convergent evolution or could be due to insertion of a new intron subsequent to their divergence from a common ancestral gene. On the basis of the structural and functional analyses, we conclude that AR is the third member of the EGF/TGF-α family of growth modulators present in the human genome.

The AR gene maps to a site of frequent chromosomal aberrations in acute leukemia. We have mapped the AR gene to chromosome region 4q13-4q21. This region also contains the genes for melanoma growth-stimulatory activity (*gro*), platelet factor 4 (PF4), the gamma interferon-inducible factor IP-10, interleukin-8, statherin (a calcium-regulating salivary protein), albumin, and α-fetoprotein (13). The gene for EGF is located distally at 4q25, approximately 30×10^6 bp away, while the genes for c-*kit* and the platelet-derived growth factor A-type receptor are located closer to the centromere. *gro*, IP-10, and PF4 belong to a class of structurally related peptides that may constitute a family of growth factors clustered on the proximal part of the long arm of chromosome 4 (56). AR shows no sequence similarity to this family of proteins. Although TGF-α shares 32 to 33% homology with AR and EGF, it is located on a separate chromosome at region 2p13 (70).

Chromosomal abnormalities are frequent in many human cancers. One-third of all cases of acute lymphoblastic leukemia involve specific translocations. The most common cytogenetic aberration in congenital acute lymphoblastic leukemia, t(4;11)(q21;q23), involves the region near the AR gene (32). ALL with t(4;11) is most common in infants under the age of 16 months, and this translocation identifies a group with a poor prognosis in need of aggressive therapy. Aberrations involving region 4q21 have also been associated with T lymphomas (43), and with the piebald trait, an inherited disorder resulting in patchy skin pigmentation due to deficient melanoblast migration and differentiation (13). Linkage studies between AR and these genetic disorders will allow us to determine the significance of their colocalization.

Regulation of AR expression. The tissue distribution of AR transcripts was distinct from that of EGF, but had some similarities with the distribution of TGF-α mRNA. EGF is expressed predominantly in the submaxillary glands, the distal tubules of the kidney, and lactating mammary gland, with lower levels in the pancreas, small intestine, and pituitary (54). TGF-α is expressed in a variety of tumors and retrovirally transformed cells (15, 17) and in early fetal development and preimplantation embryos (55, 73). Initial attempts to identify a normal adult tissue source of TGF-α expression were unsuccessful, in part because of low mRNA abundance and cell type specificity (17). TGF-α has subsequently been shown to play a role in the growth of normal cells, since its mRNA has been detected in normal and psoriatic adult keratinocytes, breast epithelial cells, pituitary, brain, ovarian thecal cells, seminiferous tubules, and activated alveolar macrophages (21, 46, 67, 73). Both TGF-α and AR are expressed in the normal ovary, testis, and breast tissue, and this paper shows that AR can also be detected in

the placenta, pancreas, heart, colon, lungs, spleen, and kidneys. This expression profile suggested that AR serves a functional role in the growth of normal tissues and that it may be involved in early development or in processes as diverse as gonadogenesis, hematopoiesis, and tissue modeling and repair.

Since AR was first discovered as a growth regulator from TPA-stimulated cells, we studied the mechanism of this induction. Several tumor cell lines showed increased AR mRNA after 24 h of treatment with TPA. This prolonged treatment with TPA was selected since induction of AR transcripts in MCF-7 cells peak at 24 h and their supernatant contains the largest amount of AR protein after a 48-h induction. On characterization of the 5′ regulatory region of the AR gene, a promoter was identified that was TPA responsive in MCF-7 cells yet lacked the consensus TPA-responsive element sequence. Similarly, TGF-α expression has been reported to be induced by TPA in MDA-MB-468 cells (4) and in human keratinocytes (53), yet its 5′-flanking region also lacks a consensus TPA-responsive element (36). Maximal induction of AR was seen following prolonged treatment with TPA, conditions sufficient to induce down regulation of protein kinase C (60). Possibly more complex interactions are required for TPA induction of AR. A potential cAMP-responsive element exists in the 5′-flanking region of AR, and recent reports demonstrate that cross-talk and synergy occur between the protein kinase C and cAMP pathways (19, 51). cAMP alone can activate the 8-bp TPA-responsive element, but TPA has not been shown to interact with the 7-bp cAMP-responsive element (19). Evidence that phorbol esters stimulate cAMP accumulation provides a possible link between these two pathways (78).

Conclusions. AR is unique among growth regulators in that it has the potential for two disparate mechanisms of signal transduction. First, like EGF, TGF-α, and other growth regulators, AR binds to a membrane receptor that phosphorylates protein(s) in the cytoplasm of target cells and thus generates a second messenger(s) that regulates gene expression in a spatial and temporal manner. Second, AR is endowed with nuclear-targeting motifs and has the capacity to interact directly with DNA; by this mechanism, gene expression could be controlled in a fashion similar to that observed with glucocorticoid, thyroid, and morphogen receptors (22).

The AR precursor may also exist as an integral membrane "receptor" and function in cell-cell interactions or in the regulation of cell growth and differentiation. Such membrane-bound forms might be involved in eliciting a program of growth regulation that is different from that elicited by the secreted proteins. An important direction for future research will be to define which residues are responsible for the differences in the activities of AR and EGF/TGF-α. One approach will be to generate chimeric and mutated forms of AR by using strategies analogous to homolog-scanning mutagenesis or by methods based on sophisticated computer modeling. These altered ligands will also allow us to further define the functional binding site by which this family of growth modulators interacts with specific cell surface receptors and transduces their biological effects.

ACKNOWLEDGMENTS

We thank Tony Purchio for his support and helpful discussions, Dave Adler for plotting histograms of the in situ hybridization data, and Jeff Murray for his generous gift of chromosome 4-specific somatic cell hybrid DNA.

LITERATURE CITED

1. **Baeuerle, P. A., and W. B. Huttner.** 1987. Tyrosine sulfation is a trans-Golgi-specific protein modification. J. Cell Biol. 105: 2655–2664.
2. **Bates, S. E., N. E. Davidson, E. M. Valverius, C. E. Freter, R. B. Dickson, J. P. Tam, J. E. Kudlow, M. E. Lippman, and D. S. Salomon.** 1988. Expression of transforming growth factor alpha and its messenger ribonucleic acid in human breast cancer: its regulation by estrogen and its possible functional significance. Mol. Endocrinol. 2:543–555.
3. **Bell, G. I., N. M. Fong, M. M. Stempien, M. A. Wormsted, D. Caput, L. L. Ku, M. S. Urdea, L. B. Rall, and R. Sanchez-Pescador.** 1986. Human epidermal growth factor precursor: cDNA sequence, expression *in vitro* and gene organization. Nucleic Acids Res. 14:8427–8446.
4. **Bjorge, J. D., A. J. Paterson, and J. E. Kudlow.** 1989. Phorbol ester or epidermal growth factor (EGF) stimulates the concurrent accumulation of mRNA for the EGF receptor and its ligand transforming growth factor-alpha in a breast cancer cell line. J. Biol. Chem. 264:4021–4027.
5. **Brachmann, R., P. B. Lindquist, M. Nagashima, W. Kohr, T. Lipari, M. Napier, and R. Derynck.** 1989. Transmembrane TGF-alpha precursors activate EGF/TGF-alpha receptors. Cell 56:691–700.
6. **Bringman, T. S., P. B. Lindquist, and R. Derynck.** 1987. Different transforming growth factor-alpha species are derived from a glycosylated and palmitoylated transmembrane precursor. Cell 48:429–440.
7. **Brown, J. P., D. R. Twardzik, H. Marquardt, and G. J. Todaro.** 1985. Vaccinia virus encodes a polypeptide homologous to epidermal growth factor and transforming growth factor. Nature (London) 313:491–492.
8. **Campbell, I. D., R. M. Cooke, M. Baron, T. S. Harvey, and M. J. Tappin.** 1989. The solution structures of epidermal growth factor and transforming growth factor alpha. Prog. Growth Factor Res. 1:13–22.
9. **Chang, W., C. Upton, S. L. Hu, A. F. Purchio, and G. McFadden.** 1987. The genome of Shope fibroma virus, a tumorigenic poxvirus, contains a growth factor gene with sequence similarity to those encoding epidermal growth factor and transforming growth factor alpha. Mol. Cell. Biol. 7:535–540.
10. **Chiu, R., M. Imagawa, R. J. Imbra, J. R. Bockoven, and M. Karin.** 1987. Multiple cis- and trans-acting elements mediate the transcriptional response to phorbol esters. Nature (London) 329:648–651.
11. **Cooke, R. M., A. J. Wilkinson, M. Baron, A. Pastore, M. J. Tappin, I. D. Campbell, H. Gregory, and B. Sheard.** 1987. The solution structure of human epidermal growth factor. Nature (London) 327:339–341.
12. **Cool, D. E., and R. T. MacGillivray.** 1987. Characterization of the human blood coagulation factor XII gene. Intron/exon gene organization and analysis of the 5′-flanking region. J. Biol. Chem. 262:13662–13673.
13. **Cox, D., and J. Murray.** 1988. Report of the committee on the genetic constitution of chromosome 4. Cytogenet. Cell Genet. 49:52–54.
14. **Defeo-Jones, D., J. Y. Tai, G. A. Vuocolo, R. J. Wegrzyn, T. L. Schofield, M. W. Riemen, and A. Oliff.** 1989. Substitution of lysine for arginine at position 42 of human transforming growth factor-alpha eliminates biological activity without changing internal disulfide bonds. Mol. Cell. Biol. 9:4083–4086.
15. **DeLarco, J. E., and G. J. Todaro.** 1978. Growth factors from murine sarcoma virus-transformed cells. Proc. Natl. Acad. Sci. USA 75:4001–4005.
16. **Derynck, R.** 1988. Transforming growth factor alpha. Cell 54:593–595.
17. **Derynck, R., D. V. Goeddel, A. Ullrich, J. U. Gutterman, R. D. Williams, T. S. Bringman, and W. H. Berger.** 1987. Synthesis of messenger RNAs for transforming growth factors alpha and beta and the epidermal growth factor receptor by human tumors. Cancer Res. 47:707–712.
18. **Derynck, R., A. B. Roberts, M. E. Winkler, E. Y. Chen, and D. V. Goeddel.** 1984. Human transforming growth factor-alpha:

precursor structure and expression in E. coli. Cell 38:287–297.
19. **Deutsch, P. J., J. P. Hoeffler, J. L. Jameson, and J. F. Habener.** 1988. Cyclic AMP and phorbol ester-stimulated transcription mediated by similar DNA elements that bind distinct proteins. Proc. Natl. Acad. Sci. USA 85:7922–7926.
20. **Di Marco, E., J. H. Pierce, T. P. Fleming, M. H. Kraus, C. J. Molloy, S. A. Aaronson, and P. P. Di Fiore.** 1989. Autocrine interaction between TGF-alpha and the EGF-receptor: quantitative requirements for induction of the malignant phenotype. Oncogene 4:831–838.
21. **Elder, J. T., G. J. Fisher, P. B. Lindquist, G. L. Bennett, M. R. Pittelkow, R. J. Coffey, Jr., L. Ellingsworth, R. Derynck, and J. J. Voorhees.** 1989. Overexpression of transforming growth factor alpha in psoriatic epidermis. Science 243:811–814.
22. **Evans, R. M.** 1988. The steroid and thyroid hormone receptor superfamily. Science 240:889–895.
23. **Feinberg, A. P., and B. Vogelstein.** 1984. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137:266–267.
24. **Gentry, L. E., and A. Lawton.** 1986. Characterization of site-specific antibodies to the erbB gene product and EGF receptor: inhibition of tyrosine kinase activity. Virology 152:421–431.
25. **Gentry, L. E., D. R. Twardzik, G. J. Lim, J. E. Ranchalis, and D. C. Lee.** 1987. Expression and characterization of transforming growth factor alpha precursor protein in transfected mammalian cells. Mol. Cell. Biol. 7:1585–1591.
26. **Goldstein, L. A., D. F. H. Zhou, L. J. Picer, C. N. Minty, R. F. Bargatze, J. F. Ding, and E. C. Butcher.** 1989. A human lymphocyte homing receptor, the Hermes antigen, is related to cartilage proteoglycan core and link proteins. Cell 56:1063–1072.
27. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. 2:1044–1051.
28. **Gray, A., T. Dull, and A. Ullrich.** 1983. Nucleotide sequence of epidermal growth factor cDNA predicts a 128,000-molecular weight protein precursor. Nature (London) 303:722–725.
29. **Gubler, U. J., and B. J. Hoffman.** 1983. A simple and very efficient method for generating cDNA libraries. Gene 25:263–269.
30. **Hanks, S. K., A. M. Quinn, and T. Hunter.** 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science 241:42–52.
31. **Heath, W. F., and R. B. Merrifield.** 1986. A synthetic approach to structure-function relationships in the murine epidermal growth factor molecule. Proc. Natl. Acad. Sci. USA 83:6367–6371.
32. **Heim, S., A. N. Bekassy, S. Garwicz, J. Heldrup, T. Wiebe, U. Kristoffersson, N. Mandahl, and F. Mitelman.** 1987. New structural chromosomal rearrangements in congenital leukemia. Leukemia 1:16–23.
33. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28: 351–359.
34. **Huttner, W. B.** 1988. Tyrosine sulfation and the secretory pathway. Annu. Rev. Physiol. 50:563–576.
35. **Idzerda, R. L., H. Huebers, C. A. Finch, and G. S. McKnight.** 1986. Rat transferrin gene expression: tissue-specific regulation by iron deficiency. Proc. Natl. Acad. Sci. USA 83:3723–3727.
36. **Jakobovits, E. B., U. Schlokat, J. L. Vannice, R. Derynck, and A. D. Levinson.** 1988. The human transforming growth factor alpha promoter directs transcription initiation from a single site in the absence of a TATA sequence. Mol. Cell. Biol. 8: 5549–5554.
37. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. 12:857–872.
38. **Krystal, G., M. Birrer, J. Way, M. Nau, E. Sausville, C. Thompson, J. Minna, and J. Battey.** 1988. Multiple mechanisms for transcriptional regulation of the *myc* gene family in small-cell lung cancer. Mol. Cell. Biol. 8:3373–3381.
39. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105–132.
40. **Lasky, L. A., M. S. Singer, T. A. Yednock, D. Dowbenko, C. Fennie, H. Rodriguez, T. Nguyen, S. Stachel, and S. D. Rosen.** 1989. Cloning of a lymphocyte homing receptor reveals a lectin domain. Cell 56:1045–1055.
41. **Lathe, R.** 1985. Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations. J. Mol. Biol. 183:1–12.
42. **Lazar, E., E. Vicenzi, E. Van Obberghen-Schilling, B. Wolff, S. Dalton, S. Watanabe, and M. B. Sporn.** 1989. Transforming growth factor alpha: an aromatic side chain at position 38 is essential for biological activity. Mol. Cell. Biol. 9:860–864.
43. **Levine, E. G., D. C. Arthur, K. J. Gaul-Peczalska, T. W. LeBien, B. A. Peterson, D. D. Hurd, and C. D. Bloomfield.** 1986. Correlations between immunological phenotype and karyotype in malignant lymphoma. Cancer Res. 46:6481–6488.
44. **Leytus, S. P., D. C. Foster, K. Kurachi, and E. W. Davie.** 1986. Gene for human factor X: a blood coagulation factor whose gene organization is essentially identical with that of factor IX and protein C. Biochemistry 25:5098–5102.
45. **Loenen, W. A., and W. J. Brammar.** 1980. A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes. Gene 10:249–259.
46. **Madtes, D. K., E. W. Raines, K. S. Sakariassen, R. K. Assoian, M. B. Sporn, G. I. Bell, and R. Ross.** 1988. Induction of transforming growth factor-alpha in activated human alveolar macrophages. Cell 53:285–293.
47. **Marquardt, H., M. W. Hunkapiller, L. E. Hood, and G. J. Todaro.** 1984. Rat transforming growth factor type 1: structure and relation to epidermal growth factor. Science 223:1079–1082.
48. **Marshall, R. D.** 1974. The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. Biochem. Soc. Symp. 40:17–26.
49. **Marth, J. D., C. Disteche, D. Pravtcheva, F. Ruddle, E. G. Krebs, and R. M. Perlmutter.** 1986. Localization of a lymphocyte-specific protein tyrosine kinase gene (lck) at a site of frequent chromosomal abnormalities in human lymphomas. Proc. Natl. Acad. Sci. USA 83:7400–7404.
50. **Montelione, G. T., M. E. Winkler, L. E. Burton, E. Rinderknecht, M. B. Sporn, and G. Wagner.** 1989. Sequence-specific 1H-NMR assignments and identification of two small antiparallel beta-sheets in the solution structure of recombinant human transforming growth factor alpha. Proc. Natl. Acad. Sci. USA 86:1519–1523.
51. **Otte, A. P., P. van Run, M. Heideveld, R. van Driel, and A. J. Durston.** 1989. Neural induction is mediated by cross-talk between the protein kinase C and cyclic AMP pathways. Cell 58:641–648.
52. **Peres, R., C. Betsholtz, B. Westermark, and C. H. Heldin.** 1987. Frequent expression of growth factors for mesenchymal cells in human mammary carcinoma cell lines. Cancer Res. 47:3425–3429.
53. **Pittelkow, M. R., P. B. Lindquist, R. T. Abraham, R. Graves-Deal, R. Derynck, and R. J. Coffey.** 1989. Induction of transforming growth factor-alpha expression in human keratinocytes by phorbol esters. J. Biol. Chem. 264:5164–5171.
54. **Rall, L. B., J. Scott, G. I. Bell, R. J. Crawford, J. D. Penschow, H. D. Niall, and J. P. Coghlan.** 1985. Mouse prepro-epidermal growth factor synthesis by the kidney and other tissues. Nature (London) 313:228–231.
55. **Rappolee, D. A., C. A. Brenner, R. Schultz, D. Mark, and Z. Werb.** 1988. Developmental expression of PDGF, TGF-alpha, and TGF-beta genes in preimplantation mouse embryos. Science 241:1823–1825.
56. **Richmond, A., E. Balentien, H. G. Thomas, G. Flaggs, D. E. Barton, J. Spiess, R. Bordoni, U. Francke, and R. Derynck.** 1988. Molecular characterization and chromosomal mapping of melanoma growth stimulatory activity, a growth factor structurally related to beta-thromboglobulin. EMBO J. 7:2025–2033.
57. **Roberts, B.** 1989. Nuclear location signal-mediated protein transport. Biochim. Biophys. Acta 1008:263–280.
58. **Rogers, S., R. Wells, and M. Rechsteiner.** 1986. Amino acid sequences common to rapidly degraded proteins: the PEST

hypothesis. Science 134:364–368.
59. **Rose, T. M., G. D. Plowman, D. B. Teplow, W. J. Dreyer, K. E. Hellstrom, and J. P. Brown.** 1986. Primary structure of the human melanoma-associated antigen p97 (melanotransferrin) deduced from the mRNA sequence. Proc. Natl. Acad. Sci. USA 83:1261–1265.
60. **Rozengurt, E.** 1986. Early signals in the mitogenic response. Science 234:161–166.
61. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
62. **Savage, C. R., Jr., T. Inagami, and S. Cohen.** 1972. The primary structure of epidermal growth factor. J. Biol. Chem. 247: 7612–7621.
63. **Schuler, G. D., and M. D. Cole.** 1988. GM-CSF and oncogene mRNA stabilities are independently regulated in trans in a mouse monocytic tumor. Cell 55:1115–1122.
64. **Shaw, G., and R. Kamen.** 1986. A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46:659–667.
65. **Shoyab, M., V. L. McDonald, J. G. Bradley, and G. J. Todaro.** 1988. Amphiregulin: a bifunctional growth-modulating glycoprotein produced by the phorbol 12-myristate 13-acetate-treated human breast adenocarcinoma cell line MCF-7. Proc. Natl. Acad. Sci. USA 85:6528–6532.
66 **Shoyab, M., G. D. Plowman, V. L. McDonald, J. G. Bradley, and G. J. Todaro.** 1989. Structure and function of human amphiregulin: a member of the epidermal growth factor family. Science 243:1074–1076.
67. **Skinner, M. K., K. Takacs, and R. J. Coffey.** 1989. Transforming growth factor-alpha gene expression and action in the seminiferous tubule: peritubular cell-Sertoli cell interactions. Endocrinology 124:845–854.
68. **Sudhof, T. C., J. L. Goldstein, M. S. Brown, and D. W. Russell.** 1985. The LDL receptor gene: a mosaic of exons shared with different proteins. Science 228:815–822.
69. **Teixido, J., R. Gilmore, D. C. Lee, and J. Massague.** 1987. Integral membrane glycoprotein properties of the prohormone pro-transforming growth factor-alpha. Nature (London) 326: 883–885.
70. **Tricoli, J. V., H. Nakai, M. G. Byers, L. B. Rall, G. I. Bell, and T. B. Shows.** 1986. The gene for human transforming growth factor alpha is on the short arm of chromosome 2. Cytogenet. Cell Genet. 42:94–98.
71. **Upton, C., J. L. Macen, and G. McFadden.** 1987. Mapping and sequencing of a gene from myxoma virus that is related to those encoding epidermal growth factor and transforming growth factor alpha. J. Virol. 61:1271–1275.
72. **von Heijne, G.** 1983. Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. 133:17–21.
73. **Wilcox, J. N., and R. Derynck.** 1988. Developmental expression of transforming growth factors alpha and beta in mouse fetus. Mol. Cell. Biol. 8:3415–3433.
74. **Wilmot, C. M., and J. M. Thornton.** 1988. Analysis and prediction of the different types of beta-turn in proteins. J. Mol. Biol. 203:221–232.
75. **Wong, S. T., L. F. Winchell, B. K. McCune, H. S. Earp, J. Teixido, J. Massague, B. Herman, and D. C. Lee.** 1989. The TGF-alpha precursor expressed on the cell surface binds to the EGF receptor on adjacent cells, leading to signal transduction. Cell 56:495–506.
76. **Yarden, Y., and A. Ullrich.** 1988. Molecular analysis of signal transduction by growth factors. Biochemistry 27:3113–3119.
77. **Yochem, J., and I. Greenwald.** 1989. *glp-1* and *lin-12*, genes implicated in distinct cell-cell interactions in C. elegans, encode similar transmembrane proteins. Cell 58:553–563.
78. **Yoshimasa, T., D. R. Sibley, M. Bouvier, R. J. Lefkowitz, and M. G. Caron.** 1987. Cross-talk between cellular signalling pathways suggested by phorbol-ester-induced adenylate cyclase phosphorylation. Nature (London) 327:67–70.
79. **Zajchowski, D., V. Band, N. Pauzie, A. Tager, M. Stampfer, and R. Sager.** 1988. Expression of growth factors and oncogenes in normal and tumor-derived human mammary epithelial cells. Cancer Res. 48:7041–7047.

had the additional scans, there were still significant increases in temporopolar blood during the production of anticipatory anxiety.
20. B. R. Kaada, *Handbook of Physiology*, Section 1, *Neurophysiology*, J. Field and H. W. Magoun, Eds. (American Physiological Society, Washington, DC, 1960), vol. 2, pp. 1345–1372.
21. W. Penfield and H. Jasper, *Epilepsy and the Functional Anatomy of the Human Brain* (Little, Brown, Boston, 1954).
22. A. Kling and H. D. Steklin, *Brain Behav. Evol.* **13**, 216 (1976).
23. E. Halgren and R. Walter, *Brain* **101**, 83 (1978); P. Gloor, A. Olivier, L. F. Quesney, F. Andermann, S. Horowitz, *Ann. Neurol.* **12**, 129 (1982); E. Strauss, A. Risser, M. W. Jones, *Arch. Neurol.* **39**, 626 (1982).
24. M. Wall *et al.*, *J. Clin. Psychiatry* **46**, 143 (1985); M. Wall, D. Mielke, J. S. Luther, *ibid.* **47**, 219 (1986); A. M. Ghadirian *et al.*, *ibid.*, p. 270.
25. M. M. Mesulam and E. J. Mufson, *J. Comp. Neurol.* **212**, 1 (1982); M. A. Moran, E. J. Mufson, M. M. Mesulam, *ibid.*, **256**, 88 (1987).
26. There was no significant difference between the left-to-right ratio of parahippocampal blood in the second and third scans ($P = 0.97$). Thus, the abnormal parahippocampal asymmetry in patients with panic disorder seems to be unrelated to a state of anticipatory anxiety.
27. J. Talairach and G. Szikla, *Atlas d'Anatomie Stereotaxique du Telencephale* (Masson, Paris, 1967).
28. We thank J. Stern and J. Price for helpful suggestions, L. Lich for technical support, and S. Sirkin and J. Woodward for secretarial assistance. Supported by National Institute of Mental Health Physician Scientist Award MH-00615 (E.M.R.); National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) Teacher Investigator Award NS-00904 (P.T.F.); NINCDS grant NS-06833; National Heart, Lung, and Blood Institute grant HL-13851; and a grant from the Upjohn company.

29 August 1988; accepted 11 November 1988

Structure and Function of Human Amphiregulin: A Member of the Epidermal Growth Factor Family

MOHAMMED SHOYAB, GREGORY D. PLOWMAN, VICKI L. MCDONALD, J. GARRETT BRADLEY, GEORGE J. TODARO

The complete amino acid sequence of amphiregulin, a bifunctional cell growth modulator, was determined. The truncated form contains 78 amino acids, whereas a larger form of amphiregulin contains six additional amino acids at the amino-terminal end. The amino-terminal half of amphiregulin is extremely hydrophilic and contains unusually high numbers of lysine, arginine, and asparagine residues. The carboxyl-terminal half of amphiregulin (residues 46 to 84) exhibits striking homology to the epidermal growth factor (EGF) family of proteins. Amphiregulin binds to the EGF receptor but not as well as EGF does. Amphiregulin fully supplants the requirement for EGF or transforming growth factor–α in murine keratinocyte growth, but it is a much weaker growth stimulator in other cell systems.

THE LIST OF PEPTIDE GROWTH REGulators has been expanding rapidly. These factors participate in various physiological and pathological conditions, such as cellular communication, growth and development, embryogenesis, immune response, hematopoiesis, cell survival and differentiation, inflammation, tissue repair and remodeling, atherosclerosis, and cancer (*1*). The isolation, characterization, and mechanism of action of regulatory factors for growth and differentiation are of current interest because of the potential use of such regulatory factors in the diagnosis, prognosis, and therapy of neoplasia and because of what these factors reveal about the basic mechanism of normal cellular proliferation and the unrestrained growth of cancer cells. We have recently reported the isolation of a novel glycoprotein termed amphiregulin (AR), which inhibits growth of A431 human epidermoid carcinoma and other human tumor cells and stimulates proliferation of human fibroblasts and other normal and tumor cells (*2*). AR was isolated from serum-free conditioned medium of MCF-7 human breast carcinoma cells that had been treated with 12-*O*-tetradecanoylphorbol-13-acetate (*2*). We now report the complete amino acid sequence of amphiregulin and compare its biological properties with those of the other members of the epidermal growth factor (EGF) family proteins.

AR was purified to homogeneity as described (*2*). The homogeneous AR was used for all the chemical and biological studies reported here. The amino acid sequence of human AR (Fig. 1) was determined by automated Edman degradation of *N*-glycanase–treated, reduced, and *S*-pyridylethylated AR (NG-SPE-AR) and of peptide fragments obtained by cleavage of NG-SPE-AR with various endopeptidases. The carboxyl-terminal analysis of NG-SPE-AR was performed with carboxypeptidase P (*Penicillium janthinellum*). The amino-terminal analysis of NG-SPE-AR revealed the presence of two sequences, one starting at residue 1, serine, and the other starting at residue 7, valine (Fig. 1). The yield of the larger form of AR was about 20% of that of the truncated form. The larger AR thus contains six additional amino acids at the amino terminal of the truncated form of AR. The larger form of AR and the truncated AR are single chain polypeptides of 84 and 78 residues, with a calculated molecular weight of 9759 and 9060, respectively (Fig. 1). Both forms of AR have a similar carboxyl-terminal sequence as determined by carboxypeptidase P cleavage (Fig. 1), and both are biologically active.

The sequence of AR was compared with all proteins in the National Biomedical Research Foundation database (release 15, containing 6796 protein sequences), Genetic Sequence Data Bank (Bolt Beranek and Newman, Los Alamos National Laboratory; release 54) and the European Molecular Biology Laboratory DNA sequence library (release 13). These computer-aided searches revealed that AR is a novel protein and a member of the EGF family. This family includes EGF (mouse, human, and rat) (*3–5*), transforming growth factor–α (TGF-α) (*6, 7*), and poxvirus growth factors [vaccinia (VGF), myxoma (MGF), and Shope fibroma (SFGF)] (*8–10*). Tissue-type plasminogen activator (*11*), the mammalian clotting factors IX and X (*12*), the low-density lipoprotein receptor (*13*), bovine protein C (*14*), human proteoglycan core protein (*15*), product of *Drosophila notch* gene (*16*), product of *lin 12* gene (*17*), the product of cell lineage–specific gene of sea urchin *Strongylocentrotus purpuratus* (*18*), cytotactin (*19*), and product of *Pfs* gene of *Plasmodium falciparum* (*20*) also contain EGF-like domains. Alignment of AR structure with the structure of EGF-like growth factors and with other members of EGF-like proteins (Fig. 2) reveals that AR, like other members of the family, contains the hallmark six essential cysteine residues, maintains conservation of cysteine residue spacing in the pattern $CX_7CX_4CX_{10}CX_1CX_8C$, and also contains some of the characteristic and conserved amino acids. AR falls between the members of the growth factor family that look like EGF and TGF-α and those that look like the poxvirus-encoded growth factors (MGF and SFGF), especially in the use of asparagine. The amino-terminal sequence of AR has some analogy with the amino-terminal sequences of the TGF-α's (*6, 7*), VGF (*8*), and MGF (*9*) in that it is rich in prolines, serines, and threonines and, like TGF-α and VGF, has potential N-linked glycosylation sites as well as the possibility for O-linked glycosylation in the region rich in serines, threonines, and prolines. Unlike MGF and SFGF, AR does not have any potential glycosylation site within the growth factor domain of the molecule. On the basis of homology with mouse EGF (*3*) and perfect alignment of six cysteine residues, one would expect the pres-

Oncogen, 3005 First Avenue, Seattle, WA 98121.

ence of three intrachain disulfide bonds in AR involving cysteine residues 46 and 59, 54 and 70, and 72 and 81.

AR is an extremely hydrophilic protein, especially the amino-terminal half of the molecule up to residue 45. A 23–amino acid stretch from residue 23 through 45 contains only five different amino acids (ten lysines, four arginines, four asparagines, three prolines, and two glycines). A tetrapeptide Arg-Lys-Lys-Lys is repeated twice (residues 26 to 29 and 40 to 43) in AR. Such sequences have been reported to serve as a nucleus targeting signal (*21*). The hydropathy profile of AR exhibits little similarity with those of the other members of the EGF family.

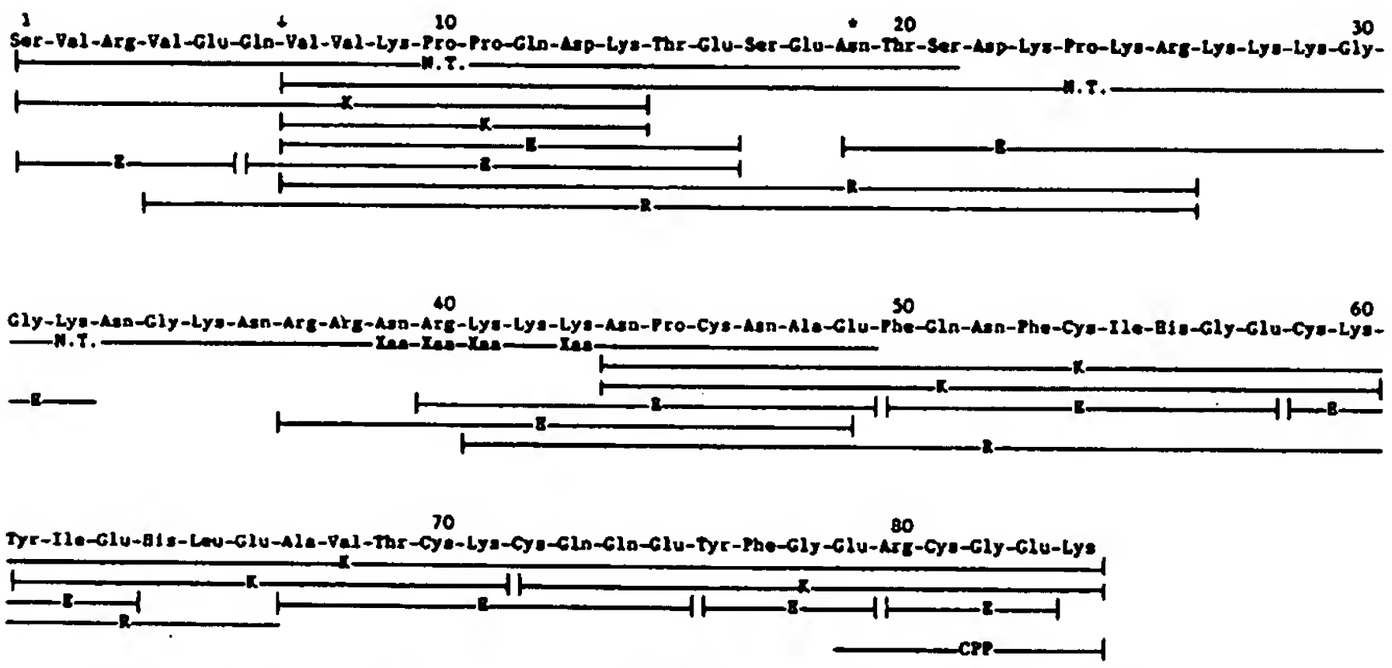


Fig. 1. Amino acid sequence of AR and schematic outline of the data supporting the sequence. The sequence of unfragmented NG-SPE-AR is denoted by N.T. Peptides obtained by cleavage with endopeptidase-Lys-C (K), with endopeptidase-Arg (R), and with endopeptidase-Glu, *Staphylococcus aureus* V8 protease (E) are indicated. CPP denotes the carboxyl-terminal sequence determined by digestion of AR with carboxypeptidase P. Residues identified with Edman degradation or by amino acid analysis are indicated by lines. Vertical bars show beginnings and endings of fragments. Lines without two vertical bars indicate incomplete sequences; ↓ indicates the start of truncated AR; and * indicates potential glycosylation site. AR was reduced with 2-mercaptoethanol and alkylated with 4-vinylpyridine. SPE-AR was purified by reversed phase high-performance liquid chromatography (rpHPLC). SPE-AR was treated with *N*-glycanase to remove N-linked oligosaccharides, and NG-SPE-AR was purified by rpHPLC. NG-SPE-AR was cleaved with various endopeptidases, and the resulting peptides were separated by an rpHPLC C_8 column. Peptide sequences were determined with an Applied Biosystems model 475A gas-phase sequencer. Identification of phenylthiohydantoin amino acid derivatives was carried out, on line, on a model 120A analyzer (Applied Biosystems). For carboxyl-terminal analysis, NG-SPE-AR was incubated with CPP, portions were withdrawn at various times, and the reaction was terminated. Released amino acids were derivatized with phenyl isothiocynate and phenylthiocarbamyl amino acid derivates were analyzed and quantitated by using micro amino acid derivatizer and analyzer (Applied Biosystem, model number 420-A0-03).

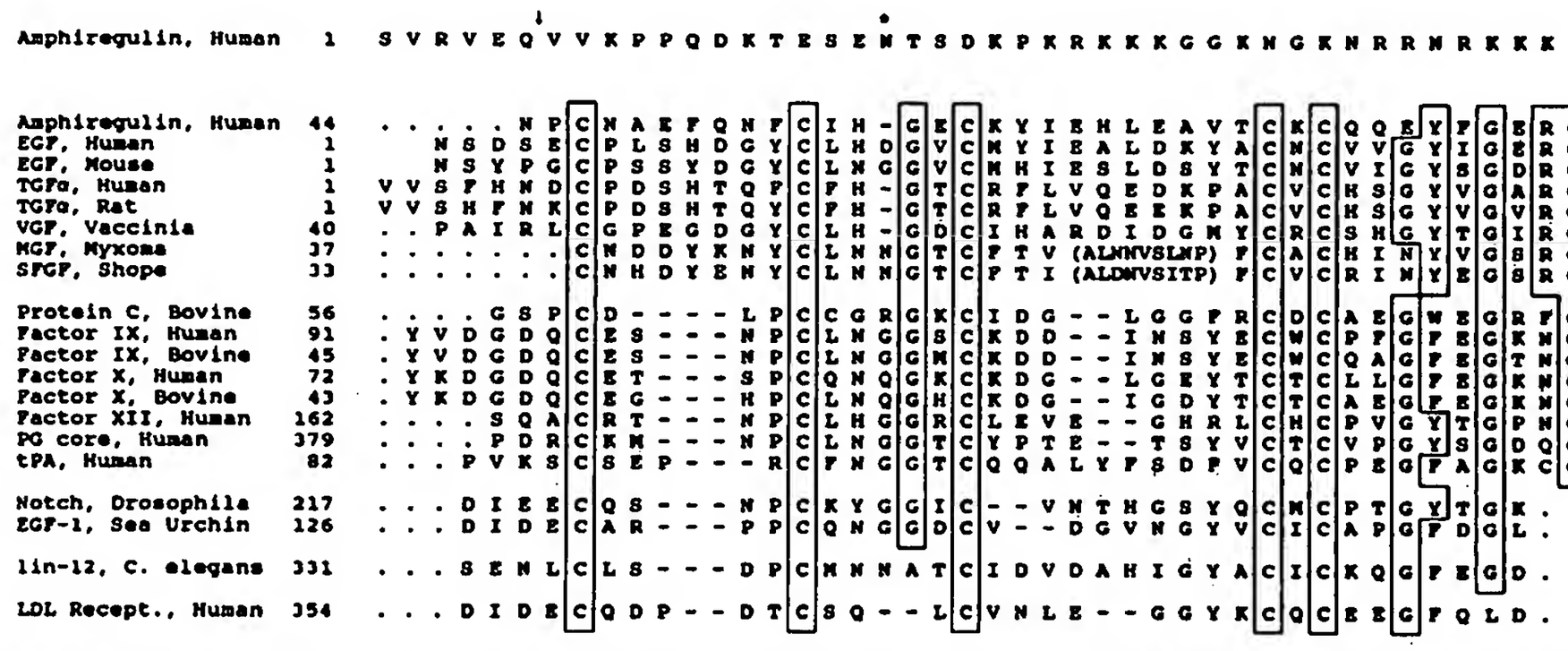


Fig. 2. Alignment of AR sequence with other EGF-like proteins. Amino acids are represented by standard one-letter symbols (*24*). Only residues appearing in eight or more proteins are boxed. Hyphens indicate gap introduced to maximize homology. Dots at the beginning and the end of sequences indicate the use of only partial sequence of a given protein. Numbers at the beginning of every sequence indicate the number of amino acid residues within the total protein sequence; ↓, beginning of the sequence of the truncated form of AR; and *, potential glycosylation site.

The binding properties of AR were compared with those of mouse EGF in radioreceptor assays (Fig. 3A). AR inhibited the binding of ^{125}I-labeled EGF to A431 cells as well as to A431 plasma membranes. A 50% inhibition of ^{125}I-EGF binding to fixed cells and membranes was seen at about 1.1 and 1.8 n*M* EGF, respectively, whereas a 50% reduction in EGF binding to cells and membranes was seen at approximately 1.8 and 5.7 n*M* AR, respectively. Unlabeled EGF completely inhibited the ^{125}I-EGF–receptor interaction at higher concentrations in both systems (Fig. 3A). However, the maximum competition with AR was 75% and 50% for binding to cells and membranes, respectively (Fig. 3A). The competition curves for AR were not parallel to that seen with EGF. These results suggest that AR has a lower affinity for EGF receptors on A431 cells than does EGF itself. Structural differences between AR and EGF might explain the binding data shown in Fig. 3. It is also possible that AR might have its specific receptor closely related to the EGF receptor.

EGF or TGF-α induce anchorage-independent growth of rat kidney cells NRK-SA6 in the presence of TGF-β (*22*). EGF induced anchorage-independent growth of NRK cells in a dose-dependent manner in the presence of TGF-β, whereas AR was found to be a noninducer of colony formation in soft agar of NRK cells (Fig. 3B). The continued growth of a murine keratinocyte cell line, Balb/MK, is dependent on EGF or TGF-α (*23*). Balb/MK cells did not proliferate in the absence of AR or EGF. However, these cells proliferated equally well in the presence of AR or of EGF (Fig. 3C). Thus,

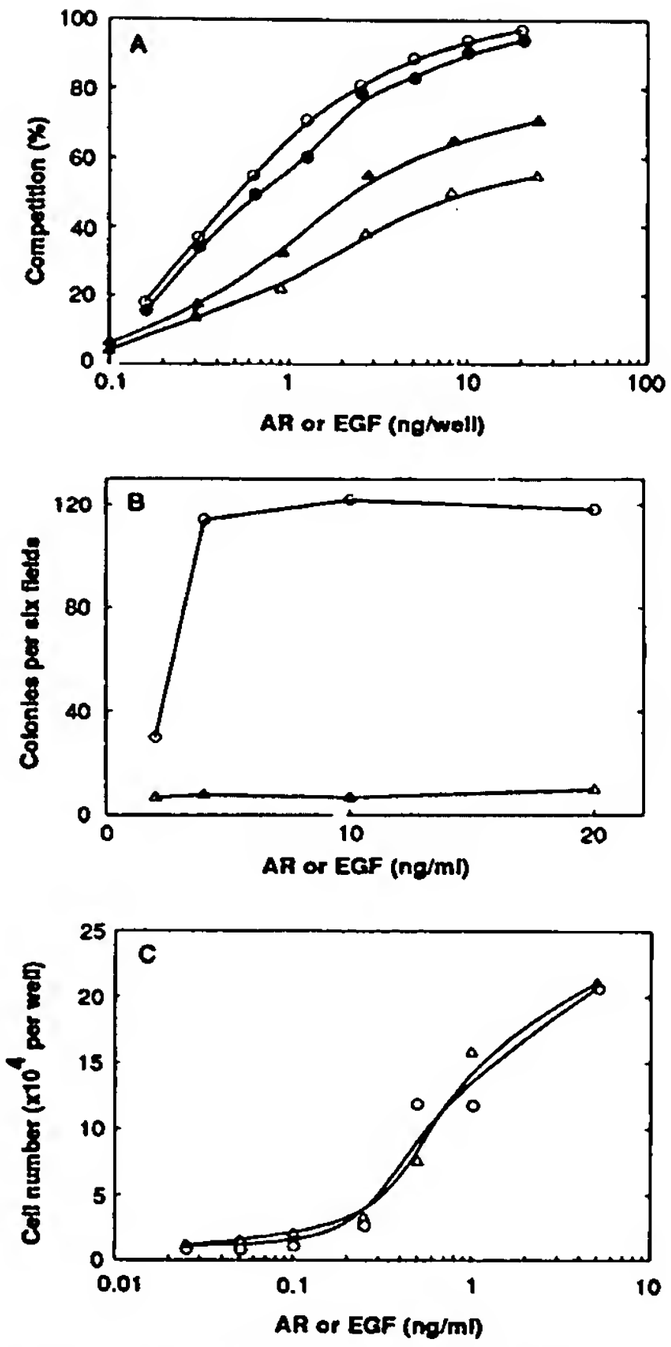


Fig. 3. (**A**) Competition of ^{125}I-EGF binding to the fixed A431 cells (solid symbols) or A431 plasma membranes (open symbols) by murine EGF and AR. EGF was radioiodinated with ^{125}I as described (*25*). The binding assays were performed either in 48-well tissue culture plates when Formalin-fixed A431 cells were used as described (*26*) or by immobilizing plasma membranes onto 96-well poly(vinyl chloride) plates as described (*27*). The binding assays used 4 ng of ^{125}I-labeled mouse EGF per milliliter, containing ~1.9 × 10^5 dpm. Samples of 100 and 50 μl were used per well for assays with fixed cells and membranes, respectively. Circles indicate EGF, and triangles indicate AR. (**B**) Effect of EGF and AR on NRK-SA6 cell colony formation in soft agar in the presence of TGF-β (1 ng/ml). A 0.38-ml base layer of 0.5% agar (Agar Noble, Difco Laboratories, Detroit) in Dulbecco's minimum essential medium containing 10% heat-inactivated fetal bovine serum (FBS) was added to 24-well Costar tissue culture plates. A 0.3% agar (0.38 ml) containing the same medium-FBS mixture, 6 × 10^3 to 12 × 10^3 test cells, and the factors to be tested were overlaid on the basal layer of agar. The plates were incubated at 37°C in the humidified atmosphere of 5% CO_2 in air. Colonies were enumerated unfixed and unstained, and the number of colonies was scored between days 7 and 10. Colonies were defined as a cluster of at least eight cells. Circles, EGF; and triangles, AR. (**C**) Effect of AR and EGF on the growth of murine keratinocytes. Balb/MK cells were plated at 1 × 10^4 cells per well in 1 ml of low calcium medium (*23*) in 24-well Costar plates (area ~2 cm^2 per well) and incubated overnight at 37°C. Then media were removed and replaced with 1 ml of medium containing various concentrations of AR or EGF in triplicate. The control wells received only medium without any AR or EGF. Plates were incubated at 37°C for 4 days, then medium was removed, wells were rinsed two times with 1 ml of phosphate-buffered saline, and the cells were detached with trypsin-EDTA and counted. Circles, EGF; and triangles, AR.

AR can supplant the EGF requirement in these cells. These results indicate that, like EGF and TGF-α, AR acts as a growth stimulator, but is much weaker on some cells (normal rat kidney) and comparable on others (murine keratinocytes).

Available structural data should allow studies on the cloning, structure, topology, expression, and regulation of amphiregulin gene in both the physiological and pathological conditions. These studies may also provide clues to design agonists and antagonists of this bifunctional growth regulator.

REFERENCES AND NOTES

1. D. Evered, J. Nugent, J. Whelan, Eds., *Ciba Found. Symp.* **116** (1985); G. Carpenter and S. Cohen, *Annu. Rev. Biochem.* **48**, 193 (1979); E. Rozengurt, *Science* **234**, 161 (1986); A. B. Pardee, *Cancer Res.* **47**, 1488 (1987); A. S. Goustin, E. B. Leof, G. D. Shipley, H. L. Moses, *ibid.* **46**, 1015 (1986); M. B. Sporn and A. B. Roberts, *J. Clin. Invest.* **78**, 329 (1986); L. Sachs, *Sci. Am.* **254**, 40 (January 1986); L. J. Old, *ibid.* **258**, 59 (May 1988); B. Beutler and A. Cerami, *N. Engl. J. Med.* **316**, 379 (1987); J. E. Groopman, *Cell* **50**, 5 (1987); M. J. Clemens and M. A. McNurlan, *Biochem. J.* **226**, 345 (1985).
2. M. Shoyab, V. L. McDonald, J. G. Bradley, G. J. Todaro, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6528 (1988).
3. C. R. Savage, Jr., T. Inagami, S. Cohen, *J. Biol. Chem.* **247**, 7612 (1972); C. R. Savage, Jr., J. H. Hash, S. Cohen, *ibid.* **248**, 7669 (1973); R. F. Doolittle, D. F. Feng, M. S. Johnson, *Nature* **307**, 558 (1984).
4. H. Gregory, *Nature* **257**, 325 (1975).
5. R. J. Simpson *et al.*, *Eur. J. Biochem.* **153**, 629 (1985).
6. G. Todaro, C. Fryling, J. DeLarco, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5258 (1980); H. Marquardt, M. W. Hunkapiller, L. E. Hood, G. J. Todaro, *Science* **223**, 1079 (1984).
7. R. Derynck *et al.*, *Cell* **38**, 287 (1984); D. Lee, T. Rose, N. Webb, G. J. Todaro, *Nature* **313**, 489 (1985).
8. M. C. Blomquist, L. C. Hunt, W. C. Barker, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7363 (1984); J. P. Brown, D. R. Twardzik, H. Marquardt, G. J. Todaro, *Nature* **313**, 491 (1985).
9. C. D. Porter and L. C. Archard, *J. Gen. Virol.* **68**, 673 (1987).
10. W. Chang, C. Upton, S.-L. Hu, A. F. Purchio, G. McFadden, *Mol. Cell. Biol.* **7**, 535 (1987).
11. D. Pennica *et al.*, *Nature* **301**, 214 (1983).
12. K. Kurachi and E. W. Davie, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6461 (1982); K. Katayama *et al.*, *ibid.* **76**, 4990 (1979); B. A. McMullen *et al.*, *Biochemistry* **22**, 2875 (1983); D. L. Enfield *et al.*, *ibid.* **19**, 659 (1980); K. Titani *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3082 (1975).
13. D. W. Russell *et al.*, *Cell* **37**, 577 (1984).
14. P. Fernlund and J. Stenflo, *J. Biol. Chem.* **257**, 12170 (1982).
15. T. Krusius, K. R. Gehlsen, E. Ruoslahti, *ibid.* **262**, 13120 (1987).
16. K. A. Wharton, K. M. Johansen, T. Xu, S. Artavanis-Tsakonas, *Cell* **43**, 567 (1985); S. Kidd, M. R. Kelley, M. W. Young, *Mol. Cell Biol.* **6**, 3095 (1986).
17. I. Greenwald, *Cell* **43**, 583 (1985).
18. D. A. Hursh, M. E. Andrews, R. A. Raff, *Science* **237**, 1487 (1987).
19. F. S. Jones *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2186 (1988).
20. D. C. Kaslow *et al.*, *Nature* **333**, 74 (1988).
21. Y. Yoneda *et al.*, *Science* **242**, 275 (1988).
22. A. B. Robert *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5339 (1981).
23. B. E. Weissman and S. A. Aaronson, *Cell* **32**, 599 (1983); R. J. Coffey *et al.*, *Cancer Res.* **48**, 1596 (1988).
24. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
25. K. Barridge, *Methods Enzymol.* **50**, 54 (1978).
26. G. Carpenter, L. King, Jr., S. Cohen, *J. Biol. Chem.* **254**, 4884 (1979); J. E. DeLarco *et al.*, *ibid.* **255**, 3685 (1980).
27. E. S. Kimball and T. C. Warren, *Biochim. Biophys. Acta* **771**, 82 (1984).

28. We thank L. Kissinger for help with the binding assays and P. Yoshida for assistance in manuscript preparation.

29 July 1988; accepted 18 November 1988